THE UNIVERSITY of York

National Institute of Economic and Social Research

CHE Centre For Health Economics

Developing New Approaches to Measuring NHS Outputs and Activity

CHE Research Paper 6

Developing new approaches to measuring NHS outputs and productivity

FINAL REPORT

8 September 2005

(revised 2 December)

Diane Dawson[*], Hugh Gravelle[**],
Mary O'Mahony[***], Andrew Street[*], Martin Weale[***]
Adriana Castelli[*], Rowena Jacobs[*], Paul Kind*, Pete Loveridge[***],
Stephen Martin[****], Philip Stevens[***], Lucy Stokes[***]

[*]Centre for Health Economics, University of York. [**]National Primary Care Research and Development Centre, Centre for Health Economics, University of York. [***]National Institute for Economic and Social Research. [****]Department of Economics and Related Studies, University of York.

Table of contents

1 Introduction 7

1.1 The research remit 7

1.2 Research delivered 8

1.3 Quality 10

1.4 Value for money and technical change 11

1.5 Structure of the report 11

2 Productivity and output measurement 14

2.1 Total factor productivity growth in private markets 14

2.2 Total factor productivity growth and quality change 16

2.3 Application of TFPG methods in the NHS 17

2.4 Value weighted NHS output index 21

2.5 Changes in marginal social values over time 22

2.6 Outcomes and attribution 24

2.7 Cost and value weights 26

3 Current practice 27

3.1 The cost weighted activity index (CWAI) 27

3.2 The 'experimental' NHS cost efficiency and service effectiveness indices 28

3.2.1 The experimental cost efficiency index 29

3.2.2 The service effectiveness growth measure 31

3.3 Pharmaceuticals and prescribing 32

4 Quality adjustment with available data 33

4.1 Activities or outputs as the unit of analysis 34

4.1.1 Activities: institutional approach 34

4.1.2 Outputs: patient-centred or disease-based approach 35

4.2 Unit of hospital output 36

4.3 Alternative sources for hospital activity 38

4.4 General practitioner consultations 41

4.5 Measures of marginal social value of outputs 41

4.5.1 Unit costs 41

4.5.2 Private sector prices 43

4.5.3 International prices 44

4.5.4 Value of health 45

4.5.5 Value of waiting time 45

4.5.6 Expert groups 46

4.6 Quality adjustment for health effects of treatment 47

4.7 Quality adjustment using long term survival 52

4.8 Quality adjustment with short term survival 54

4.8.1 Simple survival adjustment 54

4.8.2 Incorporating estimates of health effects 60

4.8.3 Life expectancy and health effects 64

4.8.4 Cost of death adjustment 65

4.8.5 In-hospital versus 30 day mortality 68

4.8.6 Conclusions: survival based quality adjustment 71

4.9 Readmissions 74

4.9.1 Readmissions as health effects 74

4.9.2 Readmissions (and clinical errors and MRSA) as a deadweight loss 80

4.10 Waiting times 81

4.10.1 Waiting time as a characteristic 82

4.10.2 Waiting time as a scaling factor 83
4.10.2.1 Discounting to start of wait 85
4.10.2.2 Discounting to date of treatment with charge for waiting 89
4.10.3 Optimal waiting times 92
4.10.4 Distribution of waiting times 92
4.10.5 Outpatient waits 96
4.10.6 Waiting time adjustment: conclusion 97

4.11 Patient satisfaction 100

4.12 Discount rate on health 103

4.13 Quality adjustment for general practice 104

4.14 Atkinson principles and quality adjustment 107

5 Experimental indices of NHS output 110

5.1 General trends, index form and data sources 110

5.2 Index form 113

5.3 Spells versus episodes 114

5.4 Survival adjustments: hospital output 115
5.4.1 Simple survival adjustment 115
5.4.2 Survival and estimated health effects adjustment 121
5.4.3 Survival adjustments with health effects and life expectancy 122

5.5 Waiting time and survival adjustments: hospital output 123
5.5.1 Effect of waiting time adjustments 125
5.5.2 Outpatient waits 130

5.6 Additional quality adjustments 131
5.6.1 Adjusting for the costs of poor treatment: readmissions and MRSA 131
5.6.2 Patient satisfaction 134

5.7 Conclusions 137

6 Specimen output index 138

6.1 Introduction 138

6.2 Data 140

6.3 Cost weighted output indices 150

6.4 Health outcome weighted output indices 155

6.5 Value weighted output index 159

6.6 Conclusion 160

7 Effects of quality adjustments on hospital and NHS output indices: summary 161

8 Labour input 165

8.1 Introduction 165

8.2 Labour input in the NHS 166

8.2.1 Volume of labour input 166

8.2.2 Quality of labour input 166

8.2.3 Data sources and volume trends 167

8.2.4 Quality adjustments based on qualifications 170

8.2.5 Quality adjustments: refinements 177

8.3 Conclusion 180

9 Experimental productivity estimates 180

9.1 Labour input and labour productivity growth 181

9.2 Intermediate and capital inputs 183

9.3 Total factor productivity growth 186

10 Improving the data 188

10.1 Outcomes data 188

10.1.1 Health outcomes 188

10.1.2 Feasibility 191

10.1.3 Cost 192

10.2 Other outcome measures: patient satisfaction 192

10.3 General practice data 193

10.3.1 GP activity 193

10.3.2 GP cost weights 194

10.3.3 General practice staff 194

10.3.4 Prescribing 194

10.4 Other primary care data 195

10.5 Inputs 196

11 Conclusions and recommendations 198

11.1 Methods 198

11.1.1 The preferred approach 198

11.1.2 Methods using existing data 199

11.2 Results 200

11.2.1 Results for the hospital sector 200

11.2.2 Other quality indicators 201

11.3 Total factor productivity growth 202

11.4 Recommendations 202

11.5 Acknowledgements 205

References 206

Annex: How should NHS output be measured? 212

Table of Notation 215

1 Introduction

1.1 The research remit

In March 2004 the Department of Health commissioned a research team from the Centre for Health Economics at the University of York and the National Institute for Economic and Social Research to develop new approaches to measuring NHS outputs and productivity. The research objectives were development of:

- A comprehensive measure of NHS outputs and productivity
- Methods to facilitate regular in-year analysis of NHS productivity
- Output measures capable of measuring efficiency and productivity at sub-national levels.

The research team was also asked to co-operate with The Atkinson Review on measurement of government output and productivity for the national accounts.

Three interim reports on this research were produced (July 2004, November 2004 and June 2005) as well as memoranda on data requirements (September 2004) and methodology (January 2005, August 2005). The work was presented for scrutiny at two workshops (7 July 2004 and 17 June 2005). The research team presented work in progress to four meetings of the NHS Outputs Steering Group (7 July 2004, 2 February 2005, 10 May 2005, 20 July 2005). This is the Final Report on the research project.

The background to the research remit referred to the Public Service Agreement (PSA) following the 2002 Spending Review that "set a 'value for money' (productivity) target of 2%". The target required information on quality improvement that had not previously been measured for the NHS as a whole. While PSA targets have changed over time, it is likely that some measure of quality improvement will continue to be required in reporting performance. Quality adjusted measures of NHS output were also required for other Department of Health purposes such as monitoring the performance of Trusts and identifying the scope for efficiency gains.

It is important to appreciate that there are significant differences between the concepts

of efficiency, value for money, productivity and productivity growth that have implications for both methods of measurement and policy relevance of the resulting indices.

- **Efficiency** is measured as the ratio of output produced with given inputs relative to the maximum feasible output.
- '**Value for money**' reflects the value individuals/society place on output relative to the costs of production. This often corresponds to a cost-benefit analysis.
- **Productivity** is the ratio of a measure of total output to a measure of total inputs.
- **Productivity growth** is the change in output relative to the change in inputs. It is often interpreted as reflecting the effect of technical change on production.

Robust measurement requires precise definition of the concept to be measured. Effective employment of these measures in pursuit of policy objectives requires selection of the appropriate measure for the issue at hand.

1.2 Research delivered

The research team has responded to the research remit by delivering the following outputs.

1. A methodology for producing a comprehensive quality adjusted index of NHS output. This is referred to as the "value weighted output index". Data necessary to estimate this index are not currently available for all NHS activities but are feasible to collect. The DH has already planned or is considering collection of the relevant data.

2. Methodologies for calculating quality adjusted NHS output indices with existing data. These are cost weighted indices that incorporate varying combinations of changes in survival, health effects, waiting times, patient satisfaction, readmissions and MRSA. We present estimates of experimental indices which examine their sensitivity to different ways of measuring waiting

times, survival, and to different assumptions about the health effect, discount rates and other parameters.

3. For the small set of hospital based treatments where there are some data on health outcomes before and after treatment, we have produced a "specimen" index that illustrates how the recommended value weighted index can be populated with data on health outcomes when they become more generally available.

4. We have suggested additional data that are feasible to collect that would not only improve future measurement of NHS output but would also be of value in managing the NHS.

5. We have constructed a new index of labour input in the NHS. It combines data from a range of sources to calculate a volume measure of total hours worked and includes an adjustment to take account of increases in the skills of the workforce.

6. Using the cost weighted quality adjusted index of outputs and inputs, we have produced provisional estimates of Labour Productivity Growth and Total Factor Productivity Growth for the period 1998/99-2003/04.

7. The methodology and data used in these indices can be applied to sub-national groups of institutions (e.g. NHS Trusts).

8. For many purposes, quality adjusted measures of output and productivity growth for particular diseases and across institutional settings will be of more value to the NHS than a comprehensive index. We indicate how, with planned changes to NHS data collection, it will be feasible to produce disease specific output and productivity indices with the methodology presented in this report.

Although key data used in our output indices, predominantly from the Hospital Episode Statistics (HES), are available on a quarterly basis, we would not recommend publication of within year estimates of output growth. The quarterly HES data are

subject to significant revision and use of quarterly index numbers could be misleading.

1.3 Quality

Central to all the work reported is a method of defining and measuring "quality".

We define the quality of treatment as the level of the characteristics valued by patients and changes in quality are measured as the rate of change of these characteristics.

Given that improving the health of patients is a primary objective of the NHS, improved health *outcomes* are likely to be the most important characteristic of treatment. In addition, the literature suggests the main impact of technical change in health care has been to improve expected health outcomes—e.g. the expected health outcomes from heart surgery or management of diabetes are better today than ten years ago. There is little data on health outcomes in the NHS and hence it has not been possible to measure quality improvement, productivity growth and technical change.

For the present the main available health outcomes data are for mortality or survival rates. This is a severe limitation on any attempt to measure the quality of output or productivity since only 3% of NHS patients die soon after treatment. There is no routine data with which to measure the improvement in health following treatment for the 97% of patients who survive. It appears that this situation may change and the NHS may start collecting data on health outcomes. In Section 4 of this report, we present the structure of output indices that should be used if and when data on health outcomes in the NHS become available. The equations could be used for a subset of patients if initially outcomes data are collected for only a limited set of procedures.

It follows from our definition of quality that the unit for measuring NHS output should be the patient treated. This makes it necessary to link the *activities* directed at treatment of a patient. For example, a patient undergoing treatment for heart disease would receive prescriptions for various drugs, attend outpatient clinics, undergo

diagnostic tests, perhaps surgery and follow-up care from a GP. At present it is not possible to identify the set of activities delivered to an NHS patient with a particular condition. The Department of Health plans to introduce a patient identifier that in future will permit analysis of the care delivered to a patient across activities, institutions and over time. For the present it is necessary to continue to use counts of activities as proxies for output. However, the indices recommended could readily be adapted to a patient-based definition of output when linked data become available.

1.4 Value for money and technical change

Recent work in the US illustrates how, with data on outcomes and an ability to link activities/inputs to patients with particular conditions, it is possible to obtain approximate disease specific measures of value for money and technical change. Cutler *et al*. (2001), for example, examine improved survival rates for patients admitted with acute myocardial infarction (AMI). By placing a monetary value on quality adjusted additional years of life expectancy and dividing by the cost of inputs used for treating this group of patients, estimates can be produced of the growth in value for money. Similar work has been done for depression, schizophrenia and cataract surgery.

The DH requested the research team to produce formulae and estimates for a comprehensive index of NHS output and productivity growth. When data become available that identify the set of inputs used to treat particular conditions and the monetary value of output, the approach we outline in Section 6 can be applied to studies of individual conditions as in the US work.

1.5 Structure of the report

For any new method of measuring NHS output and productivity to be generally accepted, it is important that the methodology be well grounded in economic theory. In Section 2 we set out the theory behind measurement of Total Factor Productivity, the issues relevant to attribution of NHS activity to improvement in health outcomes and the choice of weights necessary to sum the many NHS outputs into a single index

number.

Section 3 outlines recent and current DH practice for estimating output and productivity. This provides a baseline for comparison with the quality adjusted indices provided by the research team.

In Section 4 we set out our preferred approach to measuring quality adjusted output. This is a value weighted output index that attaches monetary values to the characteristics that measure quality. We discuss the appropriate units of output and available data. While it is feasible to collect the data necessary for a value weighted output index, the data are not currently available. In the remainder of the section we explore the possibilities for estimating a quality adjusted cost weighted output index with existing data. Quality adjusting a cost weighted index is not straightforward and we set out the assumptions required. In the absence of data on health outcomes for most NHS activity, we focus on the possibility of quality adjusting for changes in long and short term survival and for changes in waiting times. In order to illustrate the impact of including some information on health outcomes, we examine the structure of an index that includes an indicative health gain for survivors. Ordinarily improvements in survival are considered an improvement in the quality of NHS care. However, for a number of conditions, the NHS provides terminal care. We examine adjustments to the quality indicator required to deal with this issue. The appropriate method for quality adjusting a cost weighted index for changes in waiting times is not obvious. We explore several alternative methods for doing this. We conclude section 4 by comparing our approach to the recommendations of the Atkinson Review.

In Section 5 we present results for an experimental quality adjusted cost weighted output index. We show the sensitivity of the index to different ways of treating mortality rates, waiting times and choice of discount rate. We also examine the feasibility of augmenting the index with available data on patient satisfaction, readmission rates and incidence of MRSA.

The results reported in Section 5 reflect what is feasible at present when estimating a comprehensive index. However, there are a few conditions for which outcomes data are available. In Section 6 we use these data to estimate a "specimen" index. We

present results for a value weighted output index and for variants of a cost weighted index that incorporate observed health gains in the survival adjustment and allow for changes in waiting times. We also use the specimen index to illustrate the effect of substituting health outcome weights for cost weights.

In Section 7 we draw on the results from Sections 5 and 6 and present two variants of the quality adjustments, one of which is our preferred variant. We show the effects for hospital sector and for overall NHS output indices.

Section 8 is devoted to measuring labour input in the NHS. It outlines methods for calculating labour volumes and quality adjusted labour input where the latter takes account of different productivities of workers, dividing the workforce by skill group. It combines data from the NHS employment census with the rich data on worker characteristics available in the Labour Force Survey.

Section 9 brings together the output measures reported in Section 5 with the labour input measures in Section 8 to derive labour productivity estimates. Using estimates for growth in intermediate inputs and capital from a range of sources, productivity estimates are shown that account additionally for these two inputs.

In Section 10 we summarise the lessons learned in the course of this research for the availability of relevant data. We stress the importance of making better use of existing data (e.g. by record linkage and diagnosis added to prescription forms), the scope for improving output measurement with data beginning to be collected (GP consultations) and the need for the NHS to routinely collect health outcomes data.

We conclude in Section 11 with recommendations on how the Department of Health can advance work on output and productivity measurement.

2 Productivity and output measurement

2.1 Total factor productivity growth in private markets

If private markets are complete and competitive, prices reflect marginal utilities of the services to consumers and the marginal costs of providers. With some additional assumptions, the measurement and interpretation of productivity growth is then straightforward. Denote the vector of outputs from a firm at a time t as $\mathbf{x}(t)$. We index the goods by j. Let $\mathbf{z}(t)$ be the vector of n inputs (types of capital, labour and materials). $v(t)$ is a parameter which captures the state of technology at time t. The technology of the firm is described by the implicit production function

$$g\left(\mathbf{x}(t),\mathbf{z}(t),v(t)\right)=0 \tag{1}$$

Assume that the technology exhibits constant returns to scale.

Differentiating (1) with respect to time gives

$$\sum_j \frac{\partial g}{\partial x_j}\dot{x}_j+\sum_n \frac{\partial g}{\partial z_n}\dot{z}_n+\frac{\partial g_i}{\partial v}\dot{v}=0 \tag{2}$$

A profit maximising firm in a competitive market will choose $\mathbf{x}(t)$, $\mathbf{z}(t)$ to satisfy $p_j=-\theta\partial g/\partial x_j$, $w_n=\theta\partial g/\partial z_n$, where $\theta=-p_1/(\partial g/\partial x_j)$ is the Lagrange multiplier on the production constraint. We can rearrange (2) as

$$\sum_j \omega_j^y\frac{\dot{x}_j}{x_j}-\sum_n \omega_n^z\frac{\dot{z}_n}{z_n}=\frac{\theta}{px}\frac{\partial g}{\partial v}\dot{v}=\omega_1^y\left(\frac{\partial x_1}{\partial v}\frac{v}{x_1}\right)\frac{\dot{v}}{v} \tag{3}$$

where $\omega_j^y=p_jx_j\Big/\sum_j p_jx_j$, $\omega_{in}^z=w_nz_{in}\Big/\sum_j p_jx_{ij}$

and y is the value of output from the firm $y=\sum_j p_jx_j$.

The left-hand side of (3) is the rate of change of a Divisia quantity index of outputs, minus the rate of change of a Divisia quantity index of inputs. Since total factor productivity (TFP) is the ratio of an index of outputs to an index of inputs, the left hand side is also a measure of total factor productivity growth (TFPG). If production

takes place with constant returns to scale, then the total value of the product is expended on the costs of the inputs and we can replace the second term on the left hand side with the rate of change of an input index based on the cost shares $w_n z_n \Big/ \sum_n w_n z_n$

The middle and last terms in (3) are equivalent expressions for the rate of technical progress. In the last term the rate of technical progress is given as the increase in one output (x_1), holding all other outputs and inputs constant, made possible by the change in technology. Thus TFPG also measures the rate of technological progress.

Technical progress increases welfare by relaxing the production constraint on the economy. Under certain assumptions total factor productivity growth can be given a direct welfare interpretation. Thus suppose that the economy is characterised by the implicit production function $g(\mathbf{x},\mathbf{z},v) = 0$ and resources are allocated to maximise current period welfare $U(\mathbf{x},\mathbf{z})$ where $\mathbf{x}$ and $\mathbf{z}$ are vectors of outputs and inputs. The Lagrangean for the welfare problem is

$$L = U(\mathbf{x},\mathbf{z}) + \lambda g(\mathbf{x},\mathbf{z},v) \tag{4}$$

and from the envelope theorem

$$dU/dv = dL/dv = \partial L/\partial v = \lambda g_v \tag{5}$$

Hence, if U is derivable from an individualistic, non-paternal welfare function, the fact that the allocation in an economy with a complete set of competitive markets maximises some such welfare function, means that TFPG is an increasing monotonic function of the change in welfare resulting from technological change.

The simple story above takes no account of changes in the stock of capital goods used to produce consumption goods. Since what is consumed no longer equals what is produced it is more complicated to give a welfare interpretation to changes in the output index, though it is possible to do so in some cases (Sefton and Weale, forthcoming).

2.2 Total factor productivity growth and quality change

A measure of TFPG in a market sector when there is quality change can be constructed in the following manner. Let the production function for a firm or sector which produces only one type (j) of output be

$$g_j\left(x_j, q_{1j}, ..., q_{Kj}, \mathbf{z}_j, v_j\right) = 0 \tag{6}$$

Here x_j is the volume or quantity of output j (the number of units produced) and q_{kj} is the amount of outcome or characteristic k produced by consumption of one unit of output j. The vector $\mathbf{q}_j$ determines the quality of the product. At the equilibrium of a market economy the price paid for a unit of output j depends on the outcomes it produces: $p_j(\mathbf{q}_j)$, and is also a measure of quality. If the market for good j is competitive a profit maximising firm's choice of output, inputs, and outcomes will satisfy $p_j = -\theta \partial g_j / \partial x_j$, $x_j \partial p_j / \partial q_{kj} = -\theta \partial g_j / \partial q_{kj}$, and $w_n = \theta \partial g_j / \partial z_{jn}$. Totally differentiating the production function with respect to time gives

$$\frac{\partial g_j}{\partial x_j}\dot{x}_j + \sum_k \frac{\partial g_j}{\partial q_{kj}}\dot{q}_{kj} + \sum_n \frac{\partial g_j}{\partial z_{jn}}\dot{z}_{jn} + \frac{\partial g_j}{\partial v_j}\dot{v}_j = 0 \tag{7}$$

and after using the profit maximising conditions, assuming constant returns to scale to substitute total cost for the value of output in the weights on the inputs, and rearranging we get

$$\frac{\dot{x}_j}{x_j} + \sum_m \left(\frac{\partial p_j}{\partial q_{kj}}\frac{q_{kj}}{p_j}\right)\frac{\dot{q}_{kj}}{q_{kj}} - \sum_n \omega_n^z \frac{\dot{z}_{jn}}{z_{jn}} = \frac{\theta}{p_j x_j}\frac{\partial g_j}{\partial v_j}\dot{v}_j = \left(\frac{\partial y_j}{\partial v_j}\frac{v_j}{x_j}\right)\frac{\dot{v}_j}{v_j} \tag{8}$$

where $\omega_n^z = w_n z_{jn} \Big/ \sum_n w_n z_{jn}$.

Thus if we do not take account of the change in quality (the middle term in the left hand side of (8)) and merely calculate the difference between the rate of growth of the output and input indices we will not be measuring the rate of technical progress (the second and last terms). Equivalently, if we define TFPG as the difference between the rates of growth of the value of output and the cost of inputs, we will typically underestimate TFPG if we do not allow for the changing value of outputs because of improvements in quality. Consequently we need to take account of the change in the mix of outcomes (characteristics) embodied in each unit of output.

Denoting the marginal effect of outcome or characteristic k on the price of output j as $\pi_{kj} \equiv \partial p_j / \partial q_{kj}$ we can write the rate of growth of the total value of output summed across all sectors ($y = \sum_j p_j x_j = \sum_j y_j$) as

$$\frac{\dot{y}}{y} = \sum_j \frac{p_j x_j}{y} \left[\sum_k \frac{\pi_{jk} q_{kj}}{p_j} \frac{\dot{q}_{kj}}{q_{kj}} + \frac{\dot{x}_j}{x_j} \right] = \sum_j \omega_j^y \left[\sum_k \omega_k^j \frac{\dot{q}_{jk}}{q_{jk}} + \frac{\dot{x}_j}{x_j} \right] \qquad (9)$$

where $\omega_k^j = \pi_{kj} q_{kj} \Big/ \sum_k \pi_{kj} q_{kj}$

In competitive equilibrium these prices represent social values as well as costs of production. Thus, in principle the prices obtained in the competitive equilibrium enable us to calculate the rate of growth of the value of output and so derive the rate of technical progress via TFPG. We need to estimate the hedonic price functions $p_j(\mathbf{q}_j)$ which relate prices to the quality of goods (Rosen, 2002). In practice there are considerable difficulties even in market sectors in allowing for quality changes.

The discussion shows that a measure of TFPG which relates only to the volume of outputs and ignores their outcome or quality characteristics is incomplete. Note also that it is also important to capture any quality change in inputs as well as outputs. Thus if the NHS is employing more skilled labour, a measure that merely counts number of workers without taking account of differences in marginal productivities across skill types, will overestimate TFPG. The contribution from using better quality labour is incorrectly attributed to technical progress. Section 8 deals with quality adjusting labour input.

2.3 Application of TFPG methods in the NHS

The construction of a NHS productivity measure should capture the valuable things that the NHS produces. However, operationalising this simple idea is not straightforward because of the difficulties of defining NHS outputs, attaching values to the outputs, and obtaining the relevant data.

We distinguish *activities* (operative procedures, diagnostic tests, outpatient visits, consultations…), *outputs* (courses of treatment which may require a bundle of

activities), and *outcomes* (the characteristics of output which affect utility). The focus in health economics has been on the change in health produced by a course of treatment, typically measured in quality adjusted life years (QALYs). But other characteristics of treatment also affect utility: the length of time waited for treatment, the degree of uncertainty attached to the waiting time, distance and travel time to services, the interpersonal skills of GPs, the range of choice and quality of hospital food, the politeness of the practice receptionist, the degree to which patients feel involved in decisions about their treatment, etc. The aim is to measure the change in the volume of NHS outputs taking account of quality changes (changes in the volume of characteristics produced) but not of changes in the marginal social value of those characteristics. The distinction between outputs and outcomes is identical to that between goods and characteristics in consumption technology models (Deaton and Muellbauer, 1980, Ch. 10; Lancaster, 1971) where consumers value goods because of the bundle of characteristics that yield utility. The *quality* of the output is a function of the vector of outcomes it produces.

In the measurement of private sector productivity growth the focus is on outputs rather than the characteristics they produce because of the assumption that the market price of the output measures the consumers' marginal valuation of the bundle of characteristics from consuming the output. In measuring private sector productivity we also do not need to concern ourselves with counting activities because they are embodied in the outputs which are produced and sold.

The direct application of the methods used to measure TFPG in the private sector is problematic in the NHS for two main reasons. First, there is no final market for NHS outputs which makes calculation of TFPG more difficult. Second, NHS production may not be optimal, which undermines the welfare interpretation of TFPG.

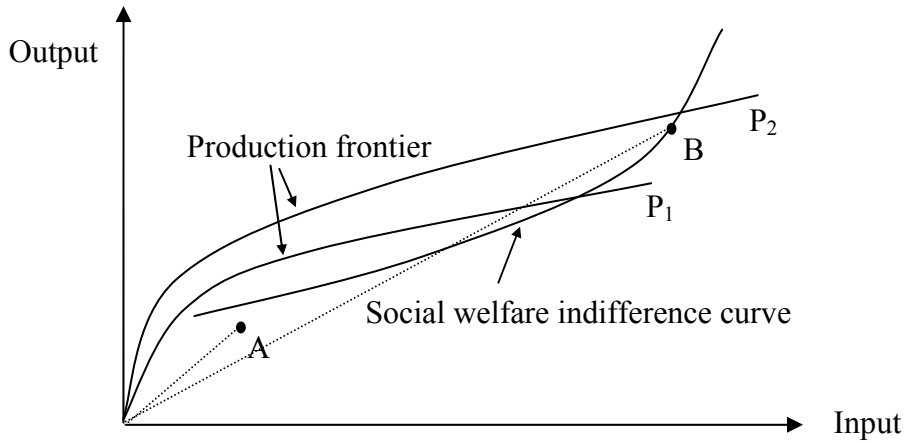
One of the justifications for having the NHS in the first place is to eliminate a market in which patients buy outputs from producers. Even in the few cases where the NHS does sell its output to the final consumer, as for pharmaceuticals prescribed by general practitioners (GPs) and dispensed to patients who are not exempt from payment, the price does not equal marginal cost.

The absence of final markets has two major consequences for attempts to measure NHS productivity. The first is that some outputs are not counted at all or are poorly measured. Instead there may be data only on the activities and even these may be lacking in many areas of activity. We discuss the implications in section 4.

The second consequence is that, because there are no prices to reveal patients' marginal valuations of NHS outputs, we have to find other means of estimating their value. We can do so in two equivalent ways: we can measure the outputs and attempt to estimate the marginal valuations attached to them or we can measure the outcomes produced by each unit of output and attempt to estimate marginal valuations of the outcomes. The bundle of outcomes produced by a unit of output is likely to change over time in the NHS because of, among other things, changes in technology or treatment thresholds. In a private market the price of output would change to reflect this. But in the absence of market prices for NHS outputs it is likely to be easier to calculate the change in the marginal value of output by focusing on the change in the vector of outcomes. We show below how the changing mix of outcomes (quality change) may be allowed for in principle. We discuss how quality adjustments based on the currently available data can be incorporated into an output index in section 4 and show the results of applying these methods to calculate experimental quality adjusted indices in Section 5. We have made suggestions as to how the quality adjustment can be improved by the collection of additional data in our Second Interim Report and discuss this further in Section 10.

The major problem in interpreting TFPG in the NHS is that it is by no means obvious that the NHS is producing optimally. It may be technically inefficient in the sense that it is possible to increase some type of output without increasing inputs or reducing some other output. It may also be producing the wrong mix of outputs. Figure 2.1 illustrates.

Figure 2.1 Productivity, efficiency and welfare



A at year 1 has higher productivity than B at year 2 but lower welfare and is less efficient (further away from its period production frontier)

Consider the simple single input, single output case in Figure 2.1. Point A in year 1 has higher productivity than point B in year 2 but welfare is lower at point A and, on any reasonable measure of technical efficiency, A has lower technical efficiency since it is further from its period production frontier. Technical progress has shifted the frontier upward from P_1 to P_2 but the productivity change does not even have the same sign as technical progress. The increase in welfare between period 1 and 2 is in part due to technical progress (B was not even feasible with the old technology) and to improvements in efficiency, perhaps because of changes in institutional structures and incentive mechanisms.

Note also that both technologies in this example have diminishing returns to scale so that increases in inputs along the frontier reduce productivity but that such a movement along the frontier can be welfare increasing.

These considerations suggest that there are problems in interpreting productivity growth as a welfare or efficiency measure. Nevertheless it can be a useful summary statistic to be used in conjunction with other data on the NHS. A further justification for attempting to measure productivity is that it will stimulate improvements in NHS information collection and processing which may lead to improved decision making within the NHS.

2.4 Value weighted NHS output index

To measure NHS TFPG we need a measure of output growth which reflects the changes in quality. Let y_{jt} be the social value of the volume of NHS output j (x_{jt}) measured at the marginal social value of output j at date t (p_{jt}). The marginal social value of output j depends on the mix of characteristics produced by a unit of output j:

$$y_{jt} = x_{jt} p_{jt} = x_{jt} \left(\sum_k \pi_{kt} q_{kjt} \right) \tag{10}$$

where q_{kjt} is the amount of characteristic k produced by a unit of j. Notice that we assume that the marginal social value function is linear

$$p_{jt} = \sum_k \pi_{kt} q_{kjt} \tag{11}$$

The assumptions that the marginal social value of a unit of output j is a linear function of its characteristics and that the π_{kt} is independent of j are strong. The latter for example requires that an improvement in the quality of hospital food (say) per day in hospital has the same effect on the value of treatment for throat cancer as on the value of a hip replacement.

The total value of NHS output is $y_t = \sum_j y_{jt}$. We want to measure the discrete time version of the growth rate of y. We could use the Tornqvist discrete time approximation to the continuous Divisia index (9) but for simplicity present the analysis in terms of a base weighted index.[1] In practice there is little difference between a chained base weighted index and the Tornqvist index. The base *value weighted output index* that we seek to measure is

$$I_{yt}^{xq} = \frac{\sum_j x_{jt+1} \sum_k \pi_{kt} q_{kjt+1}}{\sum_j x_{jt} \sum_k \pi_{kt} q_{kjt}} \tag{12}$$

which allows for changes in volume of outputs (x_{jt}) and of their characteristics (q_{kjt}) but holds the marginal value of the characteristics (π_{kt}) constant.

We can also express the value weighted index as

[1] We compare the results obtained from calculating base weighted indices with those from current weighted indices and Fisher indices (the square root of the product of the current and base weighted indices).

$$I_{yt}^{xq} = \sum \left(\frac{x_{jt+1}}{x_{jt}} \right) \left(\frac{\sum_k \pi_{kt} q_{kjt+1}}{\sum_k \pi_{kt} q_{kjt}} \right) \left(\frac{p_{jt} x_{jt}}{y_t} \right)$$

$$= \sum_j \left(1 + g_{xjt}\right) \left[\sum_k \left(1 + g_{qkjt}\right) \omega_{pjt}^{kt} \right] \omega_{yt}^{jt} \qquad (13)$$

where g_{xjt} is the growth rate of output x_j, g_{qkjt} is the growth rate of characteristic k produced by output j, ω_{pt}^{kt} is the proportion of the marginal value of output j accounted for by the k'th characteristic, and ω_{yt}^{jt} is share of the total value of period t output accounted for by output j.

Note that year to year changes in the marginal value of characteristics (π_{kt}) produced by an output j do not affect the year to year rate of growth of output j (g_{xjt}). They will however affect the weights for any index form with chained weights. Thus the overall, weighted average, rates of growth over a period of years will depend on the changes in the marginal values. This is precisely analogous to the effect of changing product prices in output indices for private sector goods and services.

2.5 Changes in marginal social values over time

In section 2.4 we specified the value of NHS output as $y = \sum_j p_j x_j$ $= \sum_j \sum_k \pi_k q_{jk} x_j$. In the rate of growth of the value of NHS output we assumed that the marginal social values of output (p_j) or of outcomes (π_k) were constant over time. If instead we had allowed changes in marginal values over time then the rate of growth of the value of NHS output would have been

$$\sum_j \left(1 + g_{xjt}\right) \left[\sum_k \left(1 + g_{\pi kt}\right) \left(1 + g_{qkjt}\right) \omega_{pjt}^{kt} \right] \omega_{yt}^{jt} - 1 \qquad (14)$$

where $g_{\pi kt}$ is the growth rate in the marginal value π_{kt} of characteristic k.

(14) depends both on changes in production conditions (the rates of growth of outcomes per unit of output and the rates of growth of outputs) but also on preferences (the rates of growth of the marginal social values of outcomes). We argue that changes in the marginal social values of outcomes between periods should not affect the growth rate between periods of the value of NHS output. The measure of output

growth is intended to measure changes in real value of output: the volume of outputs and the volumes of valuable characteristics they produce. Thus $g_{\pi kt}$ should not be included in the value weighted index of NHS output.

For example, under plausible assumptions the growth in the value of a QALY is determined by the rate of growth of income and the elasticity of marginal utility of income (Gravelle and Smith, 2001). But it is not affected by decisions within the NHS (except perhaps to a negligible extent because NHS decisions affect population health and thus the growth rate in income by improving worker productivity across the economy). We should not count changes in the marginal value of the QALY when calculating real NHS output growth.

This does not mean that changes in the value of QALYs and other outcomes have no relevance for decision making. Most decisions in the NHS have effects on outputs and outcomes over several periods – the health gain to a treated patient will typically accrue over several years. In evaluating these decisions the changing value of health should be taken into account: health changes accruing in different periods have different values. Changes in the value of health, and other characteristics, should affect decisions about the allocation of resources within and to the NHS. But they should not affect the calculation of changes in productivity between one period and the next, especially if the measure of productivity is intended to be used in part for monitoring the performance of the NHS.

Whilst we may want to exclude the growth in the marginal value of outcomes as contributing to TFPG we have to know whether and how the marginal values change over time in order to use the correct weights in calculating productivity growth.

Note that

$$\frac{\dot{p}_j}{p_j} = \sum_k \frac{\pi_k q_{jk}}{p_j}\left(\frac{\dot{\pi}_k}{\pi_k} + \frac{\dot{q}_{jk}}{q_{jk}}\right) \qquad (15)$$

which again brings out the importance of the distinction between outcomes and outputs. Even though we argue that the rate of growth of marginal social values should not be counted as part of productivity growth this does not mean that we

should remove all the rate of growth of marginal social value of *outputs* since part of $\dot{p}_j / p_j$ is due to changes in quality rather than to changing preferences.

2.6 Outcomes and attribution

The characteristics q_{kjt} are the marginal effects of the NHS. Thus for example we wish to measure the marginal effect of output x_{jt} on the health of individuals receiving this treatment, holding constant all other factors which affect health. Similarly the rate of growth of the effect of output j on characteristic k is the change in the marginal product from one period to the next due to changes in the technology (defined widely to include the way the NHS is organised). Since the other factors which affect the marginal product will also affect its rate of growth we should hold them constant in calculating the growth in the marginal product of NHS outputs from one year to the next. Although the effects of other factors (e.g. improvements in diet) on the marginal product should be excluded from the calculation of the growth rate for a particular year, they are not ignored because they affect the weights applied to the growth rates.

Parts of the national income accounting literature note that health depends on factors in addition to health service outputs (OECD, 2000; para 7.26 – 7.28). For example health depends on income, education, age and other factors exogenous to NHS activity. Hence it is argued one cannot use health outcomes to adjust outputs to take account of "quality" changes because changes in health outcome may not be attributable to health service outputs. But what we want is the *marginal* effect of output j on health. If the health production function is additively separable in health service outputs and other factors, then the marginal effect of a health service output is well defined irrespective of the level of other variables affecting health.

It is more plausible that the health production function is not additively separable so that the marginal effect of x_j on health q depends on the confounding factors. This does not present a fundamental argument against the use of outcomes. The longstanding practice of standardising mortality rates to produce a measure of population health suggests a way round the difficulty. Standardisation produces a measure of population health from which the effects of population structure (age and

gender strata) have been removed so that one can make comparisons of mortality across periods or areas without the confounding effects of demographic structure. Under certain circumstances direct standardization can identify the true differences in mortality. The assumptions required are non trivial (age and gender specific mortality can be affected only proportionately by area or period (e.g. Yule, 1934)) but direct standardisation is still useful. (The more common method of indirect standardisation which produces SMRs requires even stronger assumptions.)

Consider a simple example where health depends on a single NHS output x_1 and another variable not controlled by the NHS, for example education or income, x_2. The health production function is

$$Q_t = a_{0t} + a_{1t}x_{1t} + a_{2t}x_{2t} + a_{3t}x_{1t}x_{2t}$$

and the marginal product of health service output is

$$q_{1t} = \partial Q_t / \partial x_{1t} = a_{1t} + a_{3t}x_{2t} \qquad (16)$$

Generally we expect the effect of health service activity on health to depend on other factors ($a_{3t} \neq 0$). Hence the growth in the marginal health effect of NHS output, which is crucial for quality adjusting NHS output indices, is affected by changes in the confounding factor:

$$\frac{q_{t+1}}{q_t} = \frac{a_{1t+1} + a_{3t+1}x_{2t+1}}{a_{1t} + a_{3t}x_{2t}} \qquad (17)$$

To remove the effect of the confounding factor we can choose an arbitrary fixed level of the confounding factor in (17). If we think that the changes in the coefficient a_3t are not due to health service decisions then we should also standardize with respect to it as well:

$$\frac{q_{t+1}}{q_t} = \frac{a_{1t+1} + \bar{a}_3\bar{x}_2}{a_{1t} + \bar{a}_3\bar{x}_2} \qquad (18)$$

Obvious choices for $\bar{a}_3$ and $\bar{x}_2$ are their base period values or an average of the base period and current period values.

The health gains from treatment may increase simply because patients live longer.

Consider the example of an increase in life expectancy that is not due to developments in the NHS but reflects rising living standards, changes in diet etc. As a result an NHS treatment, such as a hip replacement, may produce a greater outcome (QALY gain) because the recipient of a hip replacement is on average alive for longer to enjoy the reduced pain and increased mobility resulting from the procedure. Thus the marginal product (the QALY gain) of the treatment is greater for reasons arising outside the health service. As far as possible, effects of changes to life expectancy which are quite independent of the procedures carried out should be kept out of calculation of the year on year growth rate in quality adjusted output of hip replacements. They should, of course, be allowed for in decisions about efficient resource allocation in the NHS but this is not the purpose of constructing an index of NHS output.

There will be some cases where the QALY gain may be partly due to improvements in the procedure and partly due to patients being "better behaved"- e.g. circulatory treatments produce more QALYs if patients do not smoke. In terms of (16) the production function is not separable and judgement will be needed about how to unravel the impacts of factors exogenous to the NHS.

2.7 Cost and value weights

By using costs to value outputs, a *cost weighted output index* can be calculated: a cost weighted sum of the growth rates of output

$$I_{ct}^{x} = \frac{\sum_j x_{jt+1} c_{jt}}{\sum_j x_{jt} c_{jt}} = \sum_j \left(\frac{x_{jt+1}}{x_{jt}} \right) \frac{x_{jt} c_{jt}}{\sum_k x_{kt} c_{kt}} = \sum_j (1 + g_{xjt}) \omega_{ct}^{jt} \qquad (19)$$

where c_{jt} is the unit (average cost) of output j (see section 3). The cost weighted index I_{ct}^{x} is equivalent to the value weighted quality adjusted index I_{yt}^{xq} only if

(a) quality change is zero for all characteristics of all outputs

(b) c_{jt} is proportional to the marginal social value of output (Dawson *et al.*, 2004a, section 2.11):

$$c_{jt} = \lambda_t p_{jt} = \lambda_t \sum_k \pi_{kt} q_{kjt} \qquad (20)$$

There is limited information on both characteristics and their marginal social value so

that attempts to estimate I_{yt}^{xq} are bound to involve compromises. Suppose that we could observe the changes in characteristics but not their marginal values (π_{kt}). How far does the assumption (20) that the output mix in the NHS maximises social value subject to budget constraint take us in estimating I_{yt}^{xq}? Using (20) in (19) gives

$$I_{yt}^{xq} = \sum_j \left(1+g_{xjt}\right)\left[\sum_m \left(1+g_{qkjt}\right)\frac{\pi_{kt}q_{jkt}}{c_{jt}}\right]\lambda_t\omega_{ct}^{jt}$$

$$= \sum_j \left(1+g_{xjt}\right)\left[\sum_m \left(1+g_{qkjt}\right)\omega_{pjt}^{kt}\right]\omega_{ct}^{jt} \qquad (21)$$

Knowledge of c_{jt} and λ_t is not sufficient to calculate the value weighted output index. We require knowledge of the relative importance of each characteristic ω_{pjt}^{kt} in determining the marginal social value of output: we need to know the marginal social values of each characteristic (π_{kt}), and the amount of each characteristic produced by each output (q_{kjt}). However if only one characteristic k is socially valuable then assumption (20) and knowledge of unit costs and the growth rate of the single valuable characteristic (k) is sufficient:

$$I_{yt}^{xq} = \sum_j \left(1+g_{xjt}\right)\left(1+g_{qkjt}\right)\omega_{ct}^{jt} \qquad (22)$$

In practice there is also imperfect information about the amount of the characteristics (q_{kjt}). Section 4 discusses the possibility of using the currently available data to quality adjust the cost weighted output index.

3 Current practice

The terminology employed by the Department of Health differs in some respects from that used in the economics literature. In our outline of current practice we use the Department of Health terminology but attempt to relate it to the economic concepts set out in Section 1.1 and used in this report.

3.1 The cost weighted activity index (CWAI)

Prior to 2004 the measure of annual NHS productivity change published by the Department of Health was based on estimating the change in a cost weighted activity

index (CWAI) less the change in NHS expenditure deflated by the index of NHS costs and prices, to generate a cost weighted efficiency index (CWEI).

CWAI was estimated using data on activity for twelve categories of Hospital and Community Health Service (HCHS) expenditure:

- Inpatient and day case episodes
- Outpatient, A&E and ward attenders
- Regular day patients
- Chiropody
- Family planning
- Screening
- District nursing
- Community psychiatric nursing
- Community learning disability nursing
- Dental episodes of care
- Ambulances

Each category of activity was weighted by its share in HCHS expenditure. There was no adjustment for improved health outcomes so that the only source of productivity improvement was an increase in the number of patients treated in hospital, ambulance trips, etc. per pound of real expenditure.

3.2 The 'experimental' NHS cost efficiency and service effectiveness indices

In 2004 the DH replaced the CWEI and developed two new 'interim' indices: an NHS cost efficiency index and a service effectiveness index. The approach was dictated by the need to respond to the Treasury's view that 'Value-for-money' should be measured in ways that permitted assessment of performance against a target of 1% p.a. improvement in cost efficiency and 1% p.a. improvement in service effectiveness. The latter was generally understood to refer to return on expenditure to improve quality.

3.2.1 The experimental cost efficiency index

The experimental cost efficiency index incorporates a change to the measurement of outputs and a change to the measurement of inputs. The change to the measurement of outputs involved replacing CWAI with an Output Index, which includes significantly more activities than CWAI and uses Reference Costs to weight different activities. The Output Index now counts over 1,700 categories of NHS activity and includes activity in primary care. The services covered are:

- Elective inpatients (over 500 activity categories)
- Non-elective inpatients (over 500 activity categories)
- Outpatients (around 300 activity categories)
- A&E (9 activity categories)
- Mental health services (30 activity categories)
- Primary care prescribing (almost 200 activity categories)
- Primary care consultations (5 activity categories)
- NHS Direct calls answered (1 activity category)
- NHS Direct online internet hits (1 activity category)
- Walk in centre visits (1 activity category)
- Ambulance journeys (1 activity category)
- General Ophthalmic Services (1 activity category)
- General Dental Services (1 activity category)
- Others including Critical care, Audiological Services, Pathology, Radiology, Chemotherapy, Renal dialysis, Community services, Bone marrow transplants & Rehabilitation (over 100 activity categories)

The coverage is not complete (Lee, 2004) and some of the omitted activities, such as the Prison Health Service, are not small; though others (Parentcraft Classes) seem unlikely to have a large impact on the index. But the extension of coverage is a very significant improvement.

Use of Reference Costs to weight Hospital and Community Health Services (HCHS) activity means that increases in more expensive treatments will have greater weight in the Output Index than increases in relatively low cost treatments. This is also true of

primary care prescribing which is measured as prescriptions issued and weighted by the cost of drugs prescribed. An increase in prescribing more expensive pharmaceuticals will have a greater effect on the Output Index than increased prescribing of less expensive drugs. The Output Index is currently used by ONS to measure NHS output in the National accounts.

Table 3.1 shows the relative weights for each main type of activity in the Output Index.

Table 3.1 Components of the NHS output index

	Cost share **2001/02**	**DH Output Index (Laspeyres)** **Growth in 2002/03 relative to 2001/02**
Electives+ day cases	12.84	5.10
Non-electives	20.64	4.92
Outpatients	10.99	4.19
Mental Health	9.56	3.62
GP & practice nurse consultations	12.44	10.27
Dentists	4.69	-0.61
Prescriptions	16.48	7.85
Accidents & Emergency	2.17	4.52
CCS	3.97	-0.28
Other	6.20	5.94
Total	100	5.36

In comparison with the previous CWEI, the experimental cost efficiency index includes a revised index of inputs in addition to the revised measurement of outputs. Since there are no measures of quality associated with the activities included in the Output Index, the DH has attempted to estimate expenditure on inputs net of expenditure intended to improve quality. Total expenditure on inputs is reduced by estimated expenditure on:

- Increases in capital charges
- Increases in Private Finance Initiative revenue expenditure

- Increases in HCHS drugs expenditure
- Increases in Information Technology expenditure
- Increases in clinical supplies expenditure
- Increases in Family Health Services drugs expenditure
- Cost of occupational enrichment
- Cost of grade enrichment
- Cost of reduced waiting times

The remaining expenditure on NHS services is deflated by the public sector price deflator to obtain an index of changes in real NHS inputs. The resulting productivity measure has been published as an index of NHS unit costs (Department of Health, 2004a, 2004b).

3.2.2 The service effectiveness growth measure

In the absence of data on quality improvement for all the activities included in the DH's new Output Index, and the need to quantify quality change for the Treasury, the DH has identified some areas where it believes that aspects of quality change can be measured and valued in monetary terms. Under consideration are:

- Reduced waiting times (outpatient, A&E, inpatient treatment)
- Reduced mortality rates for specific conditions (CHD and cancer)
- Improved patient experience

Discussion of how to value these quality improvements is still under way but possibilities include:

- Incorporating changes in mortality rates and estimates of the number of 'lives saved'. Given the age and gender of lives saved, an estimate could be made of the Quality Adjusted Life Years (QALYs) produced and valued at £30,000 per QALY. An alternative is the £1m per road death avoided used by the Department of Transport.
- Placing a value on reduced waiting times and patient experience using data from discrete choice experiments.

(Source: personal communication.)

3.3 Pharmaceuticals and prescribing

Prescriptions issued in primary care are counted as activities and therefore as outputs in the DH's new Output Index. The cost weight on this activity is total expenditure on the drugs prescribed. By contrast, in the hospital sector drugs are treated as inputs, not outputs. For hospital based activity, drugs prescribed only enter the output index as an element in the cost weight (Reference Cost) attached to an activity such as a bypass operation or dialysis: they are not counted as an activity.

The impact of the current treatment of prescribing in primary care as an output weighted by the cost of the drugs prescribed can be seen in Table 3.2. It is the movement toward prescribing more expensive drugs that contributes most to the growth in output.

Table 3.2 GP prescribing in the NHS output index (annual growth rates)

	2002/03	**2001/02**	**2000/01**
Number of prescriptions	5.44	5.41	5.01
Cost weighted prescriptions	7.85	7.52	6.28
Impact on overall index			
DH output index	5.24	4.22	1.82
DH output index excluding prescriptions	4.74	3.53	0.66

When the new NHS output index is used in estimates of productivity growth, prescription drugs are also counted as an input. ONS in their measure of productivity change present two variants for family health services drugs (net of receipts from prescription charges) employing deflators based on the average unit cost of all items and a Paasche price index for existing items. The latter is an attempt to adjust the deflator for the changing quality of drugs. These two variants lead to quite big differences, amounting to about half a percentage point per annum from 1995 to 2003 in real input growth (Lee, 2004, Hemingway, 2004). ONS rely on the Prescription Pricing Authority (PPA) and plan to consider the division by item in more detail in future revisions.

These problems would disappear in our "preferred" value weighted index of NHS output (12). Patients treated would be the unit of output weighted by health gain. Pharmaceuticals would be counted as an input. If prescribing more expensive drugs turned out to be cost effective in improving health outcomes, this would appear as a productivity increase.

There is no doubt that GPs add value through the activity of prescribing—otherwise all licensed drugs would be available over the counter. If this value added is not reflected in the assumption that the wage rate approximates the marginal product of GPs, a measure of this value added would be the correct weight for the activity of prescribing in the short-term cost weighted activity index.

4 Quality adjustment with available data

In this section we discuss how we can use available data to calculate a quality adjusted index of NHS output which corresponds as closely as possible to the ideal value weighted output index

$$I_{yt}^{xq} = \frac{\sum_j x_{jt+1} \sum_k \pi_{kt} q_{kjt+1}}{\sum_j x_{jt} \sum_k \pi_{kt} q_{kjt}} \qquad (12)$$

Calculation of (12) requires information on the outputs (x_{jt}), the outcomes q_{kjt}, and the marginal social values of the outcomes π_{kt}. Previous NHS output indices have been derived from information on outputs and have implicitly assumed that unit costs measure marginal social values. As we noted in section 2.7, even if this assumption is correct the resulting cost weighted index is not equivalent to the value weighted index unless there is no change in quality. With current information any outcome index will have to rely heavily on the assumption that unit costs measure marginal social value. Thus the main focus of the section is the extent to which it is possible to use additional existing data to calculate a quality adjusted cost weighted index. Sections 4.1 to 4.4 examine issues in the measurement of outputs (x), section 4.5 discusses sources of information on marginal social values (π), and sections 4.6 to 4.13 consider how existing data on long and short term survival, readmissions, MRSA, waiting

times and patient satisfaction can be used to proxy changes in outcomes as quality adjustments (q). The annex provides a flow chart showing the relationship of the various indices estimated in the report.

4.1 Activities or outputs as the unit of analysis

International guidance on the measurement of government output for national accounting purposes recommends distinguishing activities, outputs and outcomes. In the health service, *activities* would include operative procedures, diagnostic tests, outpatient visits, and consultations; *outputs* might comprise courses of treatment which may require a bundle of activities; and *outcomes* would be defined as the characteristics of output which affect utility.

4.1.1 Activities: institutional approach

NHS productivity measures have been based upon estimates of the number of particular types of activities (procedures, consultations etc) or the number of patients treated in various institutional settings (see section 3).

There are advantages to continuing within this framework. In instances where care for a patient with a particular condition is provided entirely within one setting, aggregation within the setting is equivalent to aggregation by patient pathway or disease group. It ensures compatibility with current NHS reporting systems and is likely to prove amenable to analysis at a disaggregated level. It can be a useful means for monitoring and managing lower level units within the NHS. Further, the approach would ensure consistency with other policy initiatives, most notably the Payment by Results reforms (Department of Health, 2002a).

The major disadvantage is that most patient cases pass through more than one institutional setting and their care requires several activities. For example, a patient who has a hip replacement will typically have been seen in general practice, in an outpatient department, treated as an inpatient in hospital and received after care treatment from her general practitioner and from personal social services. Such care patterns can lead to double counting and make problematic the valuation of output of

separate sectors contributing to joint production across sectors.

Current routine administrative data systems cannot track patients and their resource use as they move along care pathways across settings. Even within institutional settings data may not be appropriately linked. For example, whilst there are very detailed data on types and quantities of different drugs dispensed to the patients of individual general practitioners, they are not linked to the individual patient or even to diagnostic group, so it is not possible to say who got what prescriptions or for what condition.

4.1.2 Outputs: patient-centred or disease-based approach

The bulk of NHS activities or services are delivered to individual patients with the aim of improving their health. But a disease or patient pathway approach has demanding data requirements. The approach is being investigated by US researchers (Berndt *et al.*, 2002; Berndt, Busch and Frank, 2001; Cutler and Huckman, 2003; Shapiro, Shapiro and Wilcox, 2001) and, in the UK, by the Office for National Statistics. It is probably the best way forward in the long run but is not fully implementable with the types of data available in the NHS in the short to medium term. One key element required is linkage of patient records across activities and this improvement in the data is planned by the DH. Another requirement is the use of clinical teams to identify procedures and tests relevant to specific conditions and provision to update coding for procedures along clinical pathways as technology changes.

The relative advantages of the patient/disease group and institutional setting approaches depend on the degree of coverage, ease and timeliness of data collection; the dangers of double counting (for instance, where patients suffer multiple health problems); the ability to link to data on outcomes or prices; and the usefulness of the disaggregated measures (for instance, in changing behaviour).

For the short to medium term the lack of linked routine data means that the measurement of NHS output will be based predominantly on the measurement of activities rather than patients.

4.2 Unit of hospital output

The main source of data on hospital output (excluding outpatient activity) is the Hospital Episode Statistics (HES) which is derived from the cleaned returns submitted by hospital trusts.[2] There are four possible measures of hospital activity.

- Consultant episodes. The basic unit in HES is the consultant episode. Each observation records the treatments provided to a patient whilst they are under the care of a particular consultant. HES contains episodes which are unfinished at the start and end of each HES year.
- Finished consultant episodes (FCEs). A count of episodes means that an episode which spans two HES years would be counted in each year. FCEs are episodes which have finished by the end of the HES year, though they may have begun before the start of the HES year. The DH's new Output Index use finished consultant episodes (FCEs) since unit costs are derived from the Reference Costs data and these are defined for FCEs.
- Provider spells (PS). Around 8% of patients have more than one FCE during a spell in a hospital. It is possible to link episodes in the same spell to count provider spells.
- Continuous inpatient spells (CIPS). Some patients (around 1%) are transferred to another provider at the end of an episode and it is possible to link episodes across providers to yield continuous inpatient spells.

The amount of HES activity by year for FCEs, and CIPS is shown in Table 4.1. where, as they should, total FCEs exceed total CIPS. Both of these HES volume data are also always larger than those reported in the Reference Cost returns. The growth in activity (measured as the total numbers of Reference Cost hospital activities, and by HES based FCEs and CIPS) varies according to the measure employed, with all showing a larger increase in activity between 2002/03 and 2003/04. Appendix B describes our use of HES in more detail, including the construction of unit costs for spells.

[2] From 2003/4 HES data includes outpatient attendances and Accident and Emergency department activity but this had not been included in released databases at the time of producing this report.

CIPS more nearly correspond to the patient journey. CIPS capture most comprehensively the full package of inpatient care and they are less vulnerable to being miscounted if transfers among providers vary over time or if there are changes in how "being under the care of a consultant" is defined. We have therefore calculated most of our indices using CIPS, though we also report comparisons of CIPS and FCE based indices (section 5).

We recommend that future measures of hospital sector output use CIPS as the unit of outcome.

Table 4.1 Number of episodes, CIP spells from HES and number of episodes from Reference Costs

	1998/99	1999/00	2000/01	2001/02	2002/03	2003/04
HES data						
Episodes						
Electives	5491046	5577523	5573942	5485256	5664968	5815929
Non-electives	6486405	6613290	6692208	6802085	7025377	7516611
Total	11977451	12190813	12266150	12287341	12690345	13332540
Growth		**98/99-99/00**	**99/00-00/0**	**00/01-01/02**	**01/02-02/03**	**02/03-03/04**
Electives		1.57%	-0.06%	-1.59%	3.28%	2.66%
Non-electives		1.96%	1.19%	1.64%	3.28%	6.99%
Total		1.78%	0.62%	0.17%	3.28%	5.06%
CIP spells						
Electives	5427066	5487579	5479633	5386575	5578093	5736331
Non-electives	5783750	5618166	5595606	5607484	5963742	6411777
Total	11210816	11105745	11075239	10994059	11541835	12148108
Growth		**98/99-99/00**	**99/00-00/0**	**00/01-01/02**	**01/02-02/03**	**02/03-03/04**
Electives		1.12%	-0.15%	-1.69%	3.55%	2.84%
Non-electives		-2.86%	-0.40%	0.21%	6.35%	7.51%
Total		-0.94%	-0.27%	-0.73%	4.98%	5.25%
Reference Cost data						
Electives	4730410	4805812	5166244	5171867	5360406	5467913
Non-electives	5051451	5220380	5350960	5604390	5684987	6021765
Total	9781861	10026192	10517204	10776257	11045393	11489678
Growth		**98/99-99/00**	**99/00-00/0**	**00/01-01/02**	**01/02-02/03**	**02/03-03/04**
Electives		1.59%	7.50%	0.11%	3.65%	2.01%
Non-electives		3.34%	2.50%	4.74%	1.44%	5.92%
Total		2.50%	4.90%	2.46%	2.50%	4.02%

4.3 Alternative sources for hospital activity

There are two alternative sources of information about hospital activity:

- the Reference Cost returns and
- the Hospital Episode Statistics (HES).

Table 4.1 compares Reference Cost activity volumes with those for HES FCEs and CIPS.

The Reference Cost returns have been compiled annually since 1998 and have become steadily more comprehensive. Hospital activity is summarised as aggregated counts separately for elective inpatients, elective daycases and non-electives by each Healthcare Resource Group. Based on version 3.1 HRGs, the Reference Cost returns include up to 3×565 HRG categories for hospital activity (excluding "unclassified" HRGs).

Hospital activity is also available from the Hospital Episode Statistics. HES returns have been submitted by NHS providers since the late 1980s. HES contains data on every admitted patient, and comprises **individual** patient records, with information extracted directly from each patient's medical record.

HES provides different counts of activity to that recorded in the Reference Costs returns, the main reasons being the following:

- First, the HES data undergo a more thorough process of validation than the Reference Cost returns. Among other things, this validation strips out duplicate records and ensures assignment to the correct Healthcare Resource Group. The estimates of activity submitted in the Reference Cost returns are not subject to the same validation process.
- Second, HES counts all FCEs, whereas there is variable practice in what is recorded in the Reference Cost returns: sometimes all FCEs are recorded, sometimes only first FCEs are recorded. The main discrepancies between activity counts in HES and Reference Cost returns relate to activities with very long lengths of stay, including rehabilitation, mental health, bone marrow transplants, cystic fibrosis, etc. These are in HES but stripped out of Reference Cost activity.
- Third, there may be differences in how activity is apportioned to each year. HES includes all FCEs that are completed within the financial year. It is not clear how patients are counted in the Reference Cost returns when their hospital stay crosses the end of the financial year.

As well as being more thoroughly validated, HES is to be preferred to the Reference Cost return for the following reasons:

- Ideally, as explained in the previous section, we should be capturing each individual's journey through the health system. The best available measure of this the Continuous Inpatient Spell (CIPS). CIPS cannot be derived from Reference Cost returns.
- Being individual patient records, it is possible to aggregate the HES data in various ways. We aggregated the HES data into Healthcare Resource Groups, so that there is an equivalent set of activity categories as for the Reference Costs. But it is perfectly feasible to aggregate the data to other groupings, such as specialty or OPCS procedure. Moreover, HES data can be allocated easily to different HRG classifications, as the classification system is periodically revised. This flexibility in deciding activity categories is lacking in the Reference Cost returns, because these data have already been aggregated.
- We argue that NHS activities should be quality adjusted. For hospital activity, the HES data include items by which it is possible to make these adjustments, notably the waiting time prior to admission and the discharge status of the patient (from which mortality rates are derived). This information is not available in the Reference Cost returns.

For there reasons, we use HES based activity estimates rather than the Reference Cost returns for all elective inpatient, daycase and non-elective hospital activity. We use the Reference Costs database for the other sources of activity.

There are two broad HRG groups in HES not included in the reference costs – these have code T (mental health) and U (unclassified). In a spells calculation we need to include all HES activity. If we did not do so then patients whose spell included one of the omitted categories would be excluded. In addition it is also important that unclassified groups are included in a count of activities since it is likely that over time less and less activities get put into an unclassified category. It is necessary to impute unit costs to these activities. In the case of group T we used average unit cost for other Mental Health activities and for group U we used the median cost across FCEs.

4.4 General practitioner consultations

Estimates of consultation activity are derived from the consultations reported by respondents in the General Household Survey and are available by location (surgery, home, phone) and provider (GP, practice nurse (but only after 2000)). The estimate of the number of consultations per year is made by multiplying the number of reported consultations in the 14 days prior to interview by 26.

No allowance is made for seasonal factors - the date of the consultation varies across respondents and has also varied between rounds of the GHS. There have been implausibly large changes in the numbers of consultations reported in the GHS for some age-gender groups from one year to next. The GHS was also not undertaken in 1997/8 and 1999/2000 so that estimates for these years have to be interpolated. Data on consultations with practice nurses was not collected before 2000.

These deficiencies of the GHS as source of GP consultations information are widely recognised (Atkinson, 2005, pp 108-111). The DH has been investigating the use of GP record systems as a source of more accurate and detailed data. We have previously made detailed suggestions on how such data should be collected (Dawson, *et al.*, 2004b, 2004c).

New data from the QRESEARCH database derived from downloads from around 500 general practices has recently become available but too late for inclusion in this report. We have agreed to undertake an analysis of general practice consultation rates data from QRESEARCH for the DH which will examine what if any adjustments need to be made to QR consultation counts to produce an estimate of consultation activity. This report will be delivered separately in the Spring of 2006.

4.5 Measures of marginal social value of outputs

4.5.1 Unit costs

Current NHS practice, which follows the recommendation of Eurostat (2001), is to use production costs (such as the average costs as reported in the annually produced

Schedule of Reference Costs) as weights in the calculation of output indices. This implies that costs reflect the value that society places upon these activities. So cochlear implant (with a unit cost of £23,747 in the 2002/03 Reference Costs) is assumed to be 25 times more valuable than a normal delivery without complications (unit cost £921). The use of unit costs as weights reflecting the marginal social value of outputs has the support, albeit reluctant, of Hicks (1940) but as we have noted (section 2.7) it rests on the strong assumption that resources are allocated efficiently in the NHS so that unit costs are proportional to the marginal value of output produced. Even with this assumption the use of unit costs will not allow for quality changes (section 2.7) in the calculation of growth between one year and the next.

Reference Costs estimates of unit costs are based on allocations of fixed costs to HRGs with FCEs as the unit of measurement. We have investigated whether it would be possible to improve on this method of estimating activity costs by using regression analysis. The Second Interim Report (Dawson *et al.*, 2004c; section 3.7.1) describes how we attempted to estimate cost functions using a provider Trust level panel of data on activities and costs and the problems we encountered. Our subsequent estimations were no more successful. We describe these attempts in Appendix D. Apart from difficulties in trying to back compute total provider costs from the Reference Cost data on unit costs and activities, the main problem is one of degrees of freedom. There are more HRG activity types (approximately 550) than Trusts (approximately 180) so that even with observations over 6 years it is necessary to use quite high levels of aggregation of activities.

We feel that the unit costs in the Reference Costs are very unlikely to measure marginal costs, even long run marginal costs, because of the accounting procedures used to generate them. We understand that, as a result of the introduction of Practice Based Commissioning and Payment by Results, the DH is considering the production of a new set of unit costs for spells, rather than for FCEs. There is a danger that Payment by Results will encourage misreporting behaviour, with providers reporting their Reference Costs close to the tariff and being reluctant to divulge information about where their costs deviate from the tariff. If such behaviour is widespread, in future the Reference Cost database may not even approximate average, let alone, marginal costs.

There is Reference Cost data on unit costs from 1997/98. The data had patchy coverage in the early years: only 76% of activity currently recorded had unit costs assigned to them in 1997/98. Moreover there were some considerable fluctuations in unit costs for specific HRGs in the early years (Street and AbdulHussain, 2004).

We therefore decided to use the unit estimates for 1999/00 for all previous years. When there were missing unit cost data we used the estimates from the previous or following year if these were available. Some activities measured in HES have no corresponding unit costs in the Reference Costs databases and for these activities we felt it was better to retain them in the index by applying the weighted average reference cost for all other activities for that year. Having dealt with the question of duplicate entries we took great care to ensure that no other entries were dropped as a result of missing data. The general principle was that it was important not to lose any patients merely on the grounds that the records were less than complete. To this end we replaced missing data for HRG unit costs by averages for the whole. Other missing variables were again replaced by suitable population averages. For example, in order to determine the individuals' age we use the variable STARTAGE (age at start of episode) from the first episode in the spell. However, this is missing for some individuals, (e.g. in 2002/3 35,554 episodes did not have age recorded). These were replaced with the mean age for individuals of the same gender in the particular HRG and year. For those in sparsely-populated HRGs, missing values were replaced with the mean for the whole population.

4.5.2 Private sector prices

Under certain conditions the market prices for goods and services measure their marginal social value and hence can be aggregated for the construction of measures of the growth rate of output. One possible method of valuing NHS output might be to use prices from the private sector. Some NHS activities have close matches in the private sector. The main example is that some types of elective care are provided both in the private and public sectors. There are also a few private sector general practitioners. Non-emergency ambulance transport is similar to a taxi service.

In principle it might be possible either to use the private sector prices of outputs to

value NHS outputs or to estimate hedonic price functions to value characteristics or outcomes of NHS output such as waiting times and hotel services. However there are problems with attempting to use private health sector prices as measures of marginal social value of NHS outputs or outcomes:

(a) the private sector produces very little emergency care and relatively little non-elective care, roughly half of NHS activity.

(b) private sector outputs have a different mix of characteristics compared to the NHS. The health effects of treatment are probably broadly similar, but waiting times are much shorter and the quality of hotel services higher. Thus it would be necessary to attempt to estimate hedonic price functions to derive the marginal value of characteristics (π_{kjt}) rather than use the market price of the output to weight NHS outputs. Time and resource constraints meant that we did not consider this to be a feasible option for this project, though it may be worthwhile for the DH to commission scoping review to investigate the possibility.

(c) private patients are not a random sample of the population – they tend to be richer and better educated. Thus any estimated hedonic price function from the private sector may not predict the marginal valuations of characteristics for the general population.

(d) much private health care is purchased by insured individuals so that the market price of care will overstates its marginal value to the private patient.

(e) because the NHS is now encouraging commissioners (PCTs and general practices) to buy care from the private sector, prices for care to private patients will be increasingly influenced by the prices set by the NHS, which are based on Reference Costs.

Private sector prices are therefore unlikely to be useful as sources of marginal social values for most NHS outputs.

4.5.3 International prices

There is a precedent in cost benefit analysis for using world prices to value domestic output when domestic prices are absent or distorted. The rationale is that because trade could take place at world prices, they are legitimate measures of opportunity cost to the domestic economy. This option is not particularly useful in the valuation of UK health care outputs. There is not a significant world market in health care. In the

countries that do have published prices for health treatment, these tend to be administered prices subject to stringent domestic regulation or negotiation. It is highly unlikely that the relative prices observed in other countries will correspond to the relative value of NHS outputs.

These caveats notwithstanding, we did report in our Second Interim Report (Dawson, *et al.*, 2004c; section 3.7.2) whether the valuations of activity would be sensitive to the use of price information from other countries. We concluded that international prices were not likely to be useful as sources of relative marginal social valuations of NHS outputs. There were major differences in definitions of outputs so that it was not clear that similar outputs could be compared. Even when we were reasonably confident that the outputs were similar there were marked differences in the relative costs of treatments between different countries. For example the 2001/2 ratio of the costs of bilateral primary and primary hip replacement to the cost of a percutaneous transluminal coronary angioplasty (PTCA) was 2.54 in Australia and 0.94 in Italy. (The ratio in the Reference Costs database was 1.89). Other studies have also found marked differences in the input usage and hence costs for particular conditions (Baily and Garber, 1997).

4.5.4 Value of health

A value per QALY of £30,000 is believed to be compatible with the decisions made by the National Institute of Clinical Excellence, although they do not mention a value of life explicitly (Devlin and Parkin, 2004). The DH has commissioned research into the value of a QALY but its results are not yet available. We have taken £30000 as our reference value, assuming it applies for the year 2002/3 and adjusted it by the rate of growth of money GDP for other years.

4.5.5 Value of waiting time

A value weighted output index requires that we weight changes in the various characteristics by an estimate of the monetary value of each characteristic. One source of data on willingness to pay to reduce waiting times is evidence from discrete choice experiments. A recent review of the literature (Ryan, Odejar and Napper,

2004) reported that few studies addressed the issue of the monetary value of reducing waiting times for health care and contrasted this with the significantly greater volume of work on the value of time saving in transport. Of the six UK papers, only one sampled the English population. The other five were location or procedure specific. Ryan summarises the available evidence converting to 2002/03 prices. Propper's analysis of English data suggests estimated values between £36.25 and £94.19 for a one month reduction in waiting time. Hurst's study of waiting time for non-urgent rheumatology estimated values between £11.95 and £23.68 per week. Ryan points out that the Propper and Hurst studies give similar values assuming a linear additive model. A major limitation of the data available is its age. Propper's survey was undertaken in 1987. While it is possible to adjust prices for inflation, it is also likely that willingness to pay to reduce waiting time has changed over the last eighteen years.

We illustrate the effect of adopting a value weighted output index in place of a cost weighted index for a small subset of outputs where we have some health effects data (section 6). We have used the upper limit of the Propper evidence, £94.19 per month which corresponds to £3.13 per day in 2002/03 prices. This was the willingness to pay of retired individuals with above average incomes in the original survey. To explore the sensitivity of the index to price, we also use £50 per day which implies that a one month reduction in waiting is worth £1400. This is an arbitrary number which introspection suggests is likely to be at the high end of any willingness to pay for a reduction in waiting time for most elective care.

If the DH wishes to make a value weighted output index a regular part of reporting NHS performance, we recommend that new research is undertaken on social willingness to pay to reducing waiting times.

4.5.6 Expert groups

Clinical experts could provide estimates of the health effects of treatment without the need to deny cost-effective treatment to some patients for some treatments. They have been used in the UK for CABG (Williams, 1985), in the Netherlands to estimate burden of disease for 52 diagnostic groups accounting for 70% of health care costs,

and in the US for producing quality adjusted price indices for depression treatment (Berndt *et al.*, 2002).

We discussed the use of expert groups in our Second Interim Report (Dawson *et al.*, 2000c, section 3.1). We do not believe that they should be used to provide comprehensive annual updates of the estimated health effects. Such groups are costly to convene, organise and train. They would be useful for a limited set of major conditions, supplementing the regular annual snapshot before and after health data collected from patients that we recommend (section 10.1).

4.6 Quality adjustment for health effects of treatment

In the next four sections we consider how far it is possible to use existing data to quality adjust the output index for changes in the health effects of treatment, waiting times, and patient satisfaction with the process of care. We consider first what we would like to measure in principle.

We assume for the moment that the only valuable characteristic of NHS care is its effect on health status and examine how we might use data on post treatment mortality to produce a quality adjusted index of NHS output. In this and following subsections we consider various methods of using mortality information and combining it with other very limited data on the health effects of treatment, stressing the assumptions required.

As we are assuming in this section that health is the only relevant characteristic of health care we drop the subscript identifying the characteristic. Thus we use q_{jt}, π_t instead of q_{kjt}, π_{kt}. Denote the discounted sum of QALYs produced by the treatment if the patient survives treatment by

$$q^*_{jt} = \sum_s \delta^s \sigma^*_{jt}(s) \sum_\theta \rho^*_{jt}(\theta, s) h_{jt}(s) = \sum_s \delta^s \sigma^*_{jt}(s) h^*_{jt}(s) \qquad (23)$$

δ is the discount factor on QALYs. $h^*_{jt}(s)$ is the expected level of health s periods after treatment, conditional on being alive at time t:

$$h^*_{jt}(s) = \sum_\theta \rho^*_{jt}(\theta, s) h(\theta) \qquad (24)$$

$\sigma^*_{jt}(s)$ is the probability of surviving s periods given that the patient survived treatment j at date t, $h(\theta)$ is the health level from having health state θ, where θ is a vector of mental and physical health characteristics, and $\rho^*_{jt}(\theta, s)$ is probability of being in health state θ conditional on surviving s periods after treatment j at date t. To reduce notational complexity in examining the properties of the various indices we ignore the effect of age and gender on mortality, survival and the probability distribution of health states. Some HRGs are already age specific. A more disaggregated analysis is analytically straightforward by defining the output type by finer age categories and gender as well as HRG.

If the patient had not been treated their discounted sum of expected quality adjusted life years would have been

$$q^o_{jt} = \sum_s \delta^s \sigma^o_{jt}(s) \sum_\theta \rho^o_{jt}(\theta, s) h_{jt}(s) = \sum_s \delta^s \sigma^o_{jt}(s) h^o_{jt}(s) \qquad (25)$$

$h^o_j(s)$ is expected health if the patient would have survived s periods hence without receiving treatment j. $\sigma^o_{jt}(s)$ is the probability of surviving s periods if not treated. It depends on the probabilities of health status θ at s conditional on surviving without receiving treatment ($\rho^o_j(\theta, s)$) .

Setting health status when dead to zero, the expected increase in discounted QALYs from treatment j at time t is

$$q_{jt} = (1 - m_{jt}) q^*_{jt} - q^o_{jt} \qquad (26)$$

where m_{jt} is the probability of death within a short period of treatment j. This expression for the health effect of treatment is useful because it distinguishes three components of q_{jt} which are controllable by the NHS to different degrees and hence should be treated differently in calculating output growth rates attributable to the NHS.

The amount of health outcome produced per unit of output can change over time because of changes in

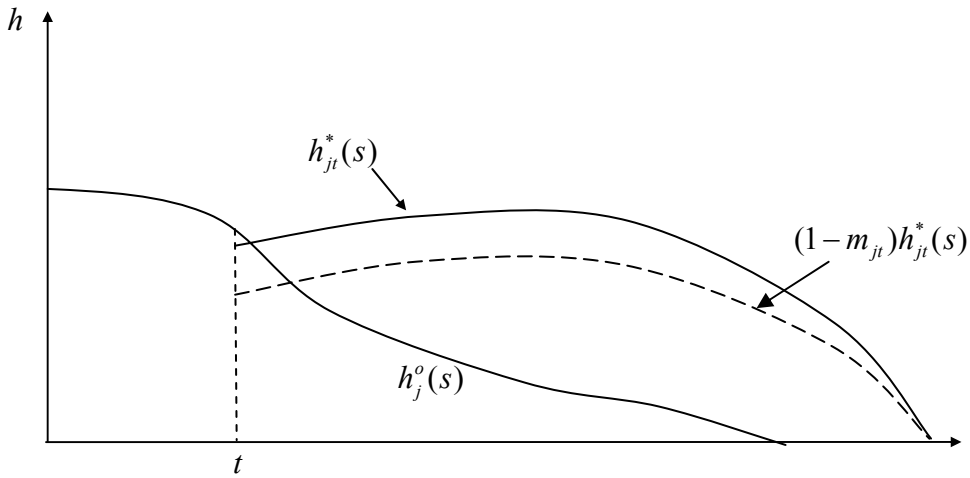
- short term post treatment mortality rate m_{jt};

- survival probabilities and health status probabilities conditional on survival $\sigma^*_{jt}(s)$, $\rho^*_{jt}(\theta, s)$
- survival and the health status probabilities conditional on not having treatment $\sigma^o_j(s)$, $\rho^o_j(\theta, s)$

The first is arguably the component most clearly attributable to the NHS for many treatments given current data and the third is unaffected by the NHS for all treatments. The effect of the NHS on the second will vary across treatments from relatively little effect on say varicose vein stripping and a large effect for cancer treatments.

Figure 4.1 illustrates the effect of treatment *j* at date *t*. The lower dashed line shows the expected time stream of health given treatment after allowing for the treatment mortality probability.

Figure 4.1 Expected time streams of health without treatment ($h^o_j(s)$), with treatment conditional on surviving treatment ($h^*_{jt}(s)$), and with treatment ($(1-m_{jt})h^*_{jt}(s)$)



Let the marginal social value at time *t* of a QALY be π_t (£s per QALY) so that the

marginal social value of unit of output *j* at time *t* is

$$p_{jt} = \pi_t q_{jt} = \pi_t \left[(1 - m_{jt}) q^*_{jt} - q^o_{jt} \right] \tag{27}$$

We wish to calculate the *value weighted output index* (12) which in the special case in which health is the only socially valuable characteristic is

$$I^{xq}_{yt} = \frac{\sum_j x_{jt+1} \pi_t q_{jt+1}}{\sum_j x_{jt} \pi_t q_{jt}} = \sum_j \left(\frac{x_{jt+1}}{x_{jt}} \frac{q_{jt+1}}{q_{jt}} \right) \frac{\pi_t q_{jt} x_{jt}}{\sum_j \pi_t q_{jt} x_{jt}} = \sum_j (1 + g_{xjt})(1 + g_{qjt}) \omega^y_{jt} \tag{28}$$

where $g_{xjt} = (x_{jt+1} - x_{jt}) / x_{jt}$ and $g_{qjt} = (q_{jt+1} - q_{jt}) / q_{jt}$ are the discrete period rates of growth of the output *j* and the health outcome per unit of output *j*.

Consider the health adjusted cost weighted output index

$$I^{xq}_{ct} = \sum_j \left(\frac{x_{jt+1}}{x_{jt}} \frac{q_{jt+1}}{q_{jt}} \frac{c_{jt} x_{jt}}{\sum_j c_{jt} x_{jt}} \right) = \frac{\sum_j x_{jt+1} \left(\frac{q_{jt+1}}{q_{jt}} \right) c_{jt}}{\sum x_{jt} c_{jt}} \tag{29}$$

The assumption of efficient allocation of NHS resources with only one quality characteristic outcome takes the form

$$c_{jt} = \lambda_t \pi_t q_{jt} = \lambda_t \pi_j \left[(1 - m_{jt}) q^*_{jt} - q^o_{jt} \right] \tag{30}$$

We can interpret $\lambda_t \pi_t$ as the cost per QALY used by the NHS in making its treatment decisions and is the optimality condition that at the margin all treatments have the same cost-effectiveness ratio. If we make the assumption (30) the health adjusted cost weighted output index is also the value weighted output index $I^{xq}_{ct} = I^{xq}_{yt}$. (This is just (22) with simpler notation.)

If we do not assume efficient allocation then we can justify calculating I^{xq}_{ct} by arguing that we are interested in a weighted average of the growth rates of the "real" parts (outputs, quality as measured by health gain per unit of output) of the value of NHS activity and that costs are convenient weights.

We can write the term q_{jt+1}/q_{jt} in (28) as

$$\frac{q_{jt+1}}{q_{jt}} = 1 + g_{qjt} = \frac{(1 - m_{jt+1}) q^*_{jt+1} - q^o_{jt+1}}{(1 - m_{jt}) q^*_{jt} - q^o_{jt}} = \frac{a_{jt+1}}{a_{jt}} \frac{q^*_{jt+1}}{q^*_{jt}} \frac{a_{jt} q^*_{jt}}{q_{jt}} - \frac{q^o_{jt+1}}{q^o_{jt}} \frac{q^o_{jt}}{q_{jt}}$$

$$= (1+g_{ajt})(1+g_{q^*_{jt}})\frac{a_{jt}q^*_{jt}}{a_{jt}q^*_{jt}-q^o_{jt}} - (1+g_{q^o_{jt}})\frac{q^o_{jt}}{a_{jt}q^*_{jt}-q^o_{jt}} \qquad (31)$$

where $a_{jt} = (1-m_{jt})$ is the survival rate (proportion of patients getting treatment j who are alive for at least a short period after treatment). g_{ajt} is the growth rate of survival and $g_{q^*_{jt}}, g_{q^o_{jt}}$ the growth rates in q^*_{jt}, q^o_{jt}.

The NHS output growth rate in any year should not include changes in q_{jt} which arise because of changes in health if not treated (q^o_{jt}). Hence in calculating the growth rate we should set $g_{q^o_{jt}} = 0$ and the health effect adjustment should be not (31) but

$$\frac{q_{jt+1}}{q_{jt}} = \left(\frac{a_{jt+1}}{a_{jt}}\right)\left(\frac{q^*_{jt+1}}{q^*_{jt}}\right)\frac{a_{jt}q^*_{jt}}{q_{jt}} - \frac{q^o_{jt}}{q_{jt}} = (1+g_{ajt})(1+g_{q^*_{jt}})\frac{a_{jt}q^*_{jt}}{q_{jt}} - \frac{q^o_{jt}}{q_{jt}} \qquad (32)$$

Notice that although we hold q^o_{jt} constant in calculating the annual growth in q_{jt} attributable to the NHS, changes in health without treatment will lead to changes in the weights.

The larger is the growth in survival (g_{ajt}) and the larger the growth in health after treatment ($g_{q^*_{jt}}$), both of which may be attributable to the NHS, the greater is the health effect adjustment factor (q_{jt+1}/q_{jt}) and the greater the index of NHS output.

In general we do not have data on health conditional on surviving treatment q^*_{jt} or conditional on no treatment q^o_{jt}. We do have data on the probability of surviving treatment a_{jt} for all hospital spells. It is also possible that in the near future we may have information on longer term survival $\sigma^*_{jt}(s)$ for a large number of NHS patients.[3] We therefore consider in the next section how it will be possible to use information on treatment survival and longer term survival.

[3] Note that we make a distinction between short term survival (*a)* and long term survival conditional on short term survival (σ^*) whereas the little data currently available is couched in terms of unconditional survival probabilities which is the product of a and σ^*.

4.7 Quality adjustment using long term survival

We want to estimate the quality adjusted cost weighted index (29) where the quality adjustment factor is given by (32). We do not know health status conditional on treatment and surviving *s* periods $h^*_{jt}(s)$. One possibility is to assume that health status *s* periods after treatment is proportional to the current health status $\bar{h}_t(s)$ of an average person who is *s* years older: $h^*_{jt}(s) = f^h_{jt}\bar{h}_t(s)$. Such data are available for example from the 1996 Health Survey for England.[4] Then we can estimate the growth in discounted expected QALYs conditional on surviving treatment as

$$\frac{q^*_{jt+1}}{q^*_{jt}} = \frac{\sum_s \delta^s \sigma^*_{jt+1}(s) f^h_{jt+1}\bar{h}_t(s)}{\sum_s \delta^s \sigma^*_{jt}(s) f^h_{jt}\bar{h}_t(s)} = \sum_s \left(\frac{\sigma^*_{jt+1}(s)}{\sigma^*_{jt}(s)}\right)\left(\frac{f^h_{jt+1}}{f^h_{jt}}\right)\left(\frac{\delta^s \sigma^*_{jt}(s) f^h_{jt}\bar{h}_t(s)}{q^*_{jt}}\right) \qquad (33)$$

Notice that we have use $\bar{h}_t(s)$ rather than $\bar{h}_{t+1}(s)$ in the numerator since changes in the general health of the population should not affect the rate of growth of health conditional on surviving treatment. We do allow for changes in the proportionality factor f^h_{jt} to affect the health conditional on short term survival since this will reflect improvements in medical technology or in patient selection, both of which should be attributed to the NHS. In the current state of knowledge we cannot generally estimate the change in the proportionality factor from one period to the next and so assume that it is constant. Hence the proportionality factors cancel from the numerator and denominator in the middle ratio in the final part of (33) and we have

$$\frac{q^*_{jt+1}}{q^*_{jt}} = \frac{\sum_s \delta^s \sigma^*_{jt+1}(s) f^h_{jt+1}\bar{h}_t(s)}{\sum_s \delta^s \sigma^*_{jt}(s) f^h_{jt}\bar{h}_t(s)} = \sum_s \left(\frac{\sigma^*_{jt+1}(s)}{\sigma^*_{jt}(s)}\right)\left(\frac{\delta^s \sigma^*_{jt}(s) f^h_{jt}\bar{h}_t(s)}{q^*_{jt}}\right) \qquad (34)$$

Thus with information on $\bar{h}_t(s), \sigma^*_{jt}(s)$ and assumptions about f^h_{jt} we can compute q^*_{jt+1}/q^*_{jt}. If it is also the case that the growth rate in survival ($g_{\sigma^* t} = \sigma^*_{jt+1}/\sigma^*_{jt}$ - 1) is constant over *s* then we do not even need to make assumptions about the magnitude of the proportionality factor: (34) simplifies to

[4] To keep the presentation simple we have assumed implicitly that all patients in an HRG have the same age and gender so that all have the same survival probabilities and expected health status. In practice when long term survival data become available it would be necessary to consider whether it was necessary to calculate age and gender specific survival rates and expected health status.

$$\frac{q^*_{jt+1}}{q^*_{jt}} = (1 + g_{q^*t}) = (1 + g_{\sigma^*t}) \tag{35}$$

We still need information on $h^o_{jt}(s), \sigma^o_{jt}(s)$ to calculate q^o_{jt} for (32) and estimates of $\overline{h}_t(s)$ are of little help since we would have to specify proportionality factors which vary across the activity types and we have no information on survival without treatment for most types. One way to proceed, which is more plausible for treatment of cancer and CHD than for cataracts, is to assume that the alternative to treatment is death so that $q^o_{jt} = 0$ and (32) becomes

$$\frac{q_{jt+1}}{q_{jt}} = \frac{a_{jt+1}}{a_{jt}} \frac{q^*_{jt+1}}{q^*_{jt}} = \frac{a_{jt+1}}{a_{jt}} \sum_s^S \left(\frac{\sigma^*_{jt+1}(s)}{\sigma^*_{jt}(s)} \right) \left(\frac{\delta^s \sigma^*_{jt}(s) f^h_t \overline{h}^*_j(s)}{q^*_{jt}} \right) \tag{36}$$

Again if we assume a constant growth in long term survival for all ages the health effect adjustment (36) simplifies further to

$$\frac{q_{jt+1}}{q_{jt}} = \frac{a_{jt+1}}{a_{jt}} \frac{q^*_{jt+1}}{q^*_{jt}} = (1 + g_{at})(1 + g_{\sigma^*t}) \tag{37}$$

There are two practical issues to consider. First to which HRGs should the adjustment in respect of say cancer survival be applied? Cancer diagnoses appear in a number of HRGs. One possibility is to apply the adjustment to the HRGs which treat the highest proportions of patients with cancer diagnoses. Alternatively one could attach group cancer patients by the HRG of their activity.

Second, we must choose a time horizon S for the adjustment. Lakhani *et al*. (2005) have recently presented estimates of 5 year cancer survival rates. The longer the horizon over which the summation in (36) takes place the more accurate the estimate of health effects. But a long time horizon has disadvantages. If the adjustment to a year is based on the survival experience of patients actually treated in that year then the output indices for S previous years will have to be revised every year. The alternative is to adjust the index for a particular year using the survival experience of patients treated S years previously. Thus the longer is S the greater the extent of revisions or the more out of date the adjustment.

Rather than use general population estimates of $\bar{h}_t(s)$ it may in some cases be reasonable to make an even cruder assumption, that survival after S years is very low or that $h^*_{jt}(s)$ after S years is very low. Setting $h^*_{jt}(s) = 0$ for $s > S$ and assuming that each year to S has the same QALY score ($h^*_{jt}(s) = h^*_j$) we get

$$\frac{q_{jt+1}}{q_{jt}} = \frac{a_{jt+1}}{a_{jt}} \frac{q^*_{jt+1}}{q^*_{jt}} = \frac{a_{jt+1}}{a_{jt}} \sum_{s=1}^{S} \left(\frac{\sigma^*_{jt+1}(s)}{\sigma^*_{jt}(s)} \right) \left(\frac{\sigma^*_{jt}\delta^s}{\sum_{s=1}^{5} \sigma^*_{jt}(s)\delta^s} \right) \tag{38}$$

If the trend improvement in survival was reasonably stable over long periods then the use of the lagged survival change data would be a reasonably accurate estimate of the adjustment based on actual survival experience since one is interested in the growth rate in survival, not in actual levels.

We believe that the use of longer term survival data is a promising way forward which will become feasible in the medium term (Lakhani *et al*., 2005). It will be especially promising if it is coupled with a programme to measure the health status of samples of NHS patients before and after treatment (see section 6; Appendix C).

4.8 Quality adjustment with short term survival

4.8.1 Simple survival adjustment

In the absence of longer term survival data we now consider what can be done to quality adjust the output index using the data which is currently available.

In the absence of information on longer term survival $\sigma^*_{jt}(s)$ we cannot estimate q^*_{jt}, the change in the discounted QALYs associated with treatment conditional on survival (q^*_{jt}). If we assume that $q^*_{jt} = q^*_{jt+1}$ does not change over time ($g_{q^*_{jt}} = 0$), so that the only reason why q_{jt} changes over time is that the post operative survival rate changes, the health effect adjustment becomes

$$\frac{q_{jt+1}}{q_{jt}} = \frac{a_{jt+1}q^*_{jt} - q^o_{jt}}{a_{jt}q^*_{jt} - q^o_{jt}} = \frac{a_{jt+1} - k_{jt}}{a_{jt} - k_{jt}} \tag{39}$$

where $k_{jt} = q^o_{jt} / q^*_{jt}$. Clearly increases in a_{jt+1} other things equal lead to a higher

quality adjustment factor. For activities with the same a_{jt} and k_{jt}, the larger the survival in period t+1 the greater the health adjustment factor.

Comparisons of quality adjustment factors across activities with different $k_{jt} = q^o_{jt} / q^*_{jt}$ require a little more thought. Remember that we are interested in the effect on the quality adjustment factor q_{jt+1}/q_{jt} which is the ratio of health effects, not in the effect of k_{jt} on the level of health effects. Suppose for definiteness that survival has increased, so that the quality adjustment factor has increased. Differentiating (39) with respect to k_{jt} gives

$$\frac{\partial(q_{jt+1}/q_{jt})}{\partial k_{jt}} = \frac{a_{jt+1} - a_{jt}}{\left(a_{jt} - k_{jt}\right)^2} \tag{40}$$

Thus outputs with a larger k_{jt} have a larger health adjustment. Higher $k_{jt} = q^o_{jt} / q^*_{jt}$ can arise from a smaller q^*_{jt} for the same q^o_{jt} or a larger q^o_{jt} for the same q^*_{jt}. The marginal effects of q^*_{jt} and q^o_{jt} are

$$\frac{\partial(q_{jt+1}/q_{jt})}{\partial q^*_{jt}} = \frac{-(a_{jt+1} - a_{jt})q^o_{jt}}{\left(a_{jt} - k_{jt}\right)^2} \tag{41}$$

$$\frac{\partial(q_{jt+1}/q_{jt})}{\partial q^o_{jt}} = \frac{(a_{jt+1} - a_{jt})q^*_{jt}}{\left(a_{jt} - k_{jt}\right)^2} \tag{42}$$

Thus, other things equal, activities with larger q^*_{jt} have smaller health quality adjustment factors. This might appear paradoxical for two reasons. First, the larger is q^*_{jt} the greater the health effect in each period so that the health adjustment factor is smaller for more beneficial activities. Second, the larger is q^*_{jt} the greater is the absolute increase $(a_{jt+1} - a_{jt})q^*_{jt}$ in the health effect between the two periods. But both "paradoxes" disappear when we remember that what we are interested in is the health adjustment factor which is the *ratio* of the health effect in the two periods. An increase in q^*_{jt} increases both numerator health effect q_{jt+1} and denominator health effect q_{jt} but has a smaller proportionate effect on q_{jt+1} than on q_{jt} and so the ratio

q_{jt+1}/q_{jt} gets smaller.

Even with the assumptions that $q^*_{jt} = q^*_{jt+1}$ is constant over short time intervals we cannot measure (39) directly unless we know the magnitudes of q^*_{jt}, q^o_{jt} in order to calculate $k_{jt} = q^o_{jt} / q^*_{jt}$. In future it may be possible to estimate q^*_{jt}, q^o_{jt} using new data on longer term survival and on health status from surveys of patients before and after treatment and from the results of evaluations of different types of treatment. But for the moment, for the vast majority of activities we have no data on q^*_{jt}, q^o_{jt}, though we do have information on survival $a_{jt.}$ We can therefore calculate the *survival adjusted cost weighted output index*

$$I^{xa}_{ct} = \frac{\sum_j c_{jt} x_{jt+1} \left(\frac{a_{jt+1}}{a_{jt}} \right)}{\sum_j c_{jt} x_{jt}} = \sum_j (1 + g_{xjt})(1 + g_{ajt}) \omega^{jt}_{ct} \qquad (43)$$

The effect of this simple survival adjustment[5] on the rate of growth of NHS productivity will not be great since the vast majority of NHS patients survive their treatment so that survival rate does not change rapidly. Thus, for example, the 30 day CIPS mortality rate for Phakoemulsion Cataract Extraction with Lens Implant (HRG B02) fell from 0.0017 in 1999/2000 to 0.0013 in 2002/3, an annual rate of decline in the mortality probability of 0.56%. The survival rate rose from 99.83 to 98.87, an annual rate of increase in the survival rate of 0.013%. In terms of Figure 4.1 the effect of increased survival is shown in the shift upward in the dashed line plotting health post treatment $(1 - m_{jt+1})h^*_j - h^o_j$. The effect is small relative to initial health.

---

[5] An alternative apparently simpler adjustment is to apply the survival rates in each period to scale the output of that period: $\sum_j c_{jt} x_{jt+1} a_{jt+1} / \sum_j c_{jt} x_{jt} a_{jt}$. Unfortunately this index equals the value weighted index $\sum_j x_{jt+1} q_{jt+1} \pi_t / \sum_j x_{jt} q_{jt} \pi_t$ under the assumptions that $q^*_{jt+1} = q^*_{jt}$, $q^o_{jt+1} = q^o_{jt} = 0$, and $c_{jt} = \theta \pi_t q^*_{jt}$. The last assumption is perverse. It is not an efficiency assumption: it requires that decision makers ignore the possibility of that a patient may not survive treatment when allocating resources across treatments. By contrast, the first two requirements and the efficiency assumption $c_{jt} = \lambda_t \pi_t \left(a_{jt} q^*_{jt} - q^o_{jt} \right)$ imply that the survival adjusted cost weighted index (43) does equal the value weighted index.

The difference between what we would like to measure (true health adjustment) and what we can measure using survival data only is

$$\frac{a_{jt+1}q^{*}_{jt+1} - q^{o}_{jt}}{a_{jt}q^{*}_{jt} - q^{o}_{jt}} - \frac{a_{jt+1}}{a_{jt}} = \frac{a_{jt}a_{jt+1}(q^{*}_{jt+1} - q^{*}_{jt}) + q^{o}_{jt}(a_{jt+1} - a_{jt})}{a_{jt}q_{jt}} \qquad (44)$$

In general we cannot say anything about the direction of the bias in using a_{jt+1} / a_{jt} instead of (32). But if survival increases and health conditional on survival increases then (44) will be positive and the simple survival adjustment will underestimate the true adjustment. If there is little change in health conditional on surviving treatment, (44) becomes

$$\frac{a_{jt+1}q^{*}_{jt} - q^{o}_{jt}}{a_{jt}q^{*}_{jt} - q^{o}_{jt}} - \frac{a_{jt+1}}{a_{jt}} = \frac{q^{o}_{jt}(a_{jt+1} - a_{jt})}{a_{jt}q_{jt}} \qquad (45)$$

and a_{jt+1} / a_{jt} will always have the same sign as (32) and will always be less than it in absolute value. a_{jt+1} / a_{jt} is a conservative estimate of the true health adjustment (32) if health conditional on surviving treatment is constant or increasing.

Table 4.2 gives some indication of the underestimation of the growth rate of the true health effect $\left(\left(a_{jt+1} - k_{jt}\right)/\left(a_{jt} - k_{jt}\right) - 1\right)$ when it is calculated as the growth rate of survival $((a_{jt+1}/a_{jt}) - 1)$. The example has a rate of survival of 0.97 in the base year which is approximately the average survival rate of patients. The greater the reduction in mortality the greater the increase in the survival rate and the greater the growth rate in the true health effect. Notice that because survival is initially high even quite large proportionate reductions in the mortality risk have small effects on the survival rate and on the true growth in the health effect. The true growth in the health effect is larger the larger is $k = q^{o}_{t} / q^{*}_{t}$. Thus, as we discussed above, the smaller the proportionate effect of treatment on the discounted sum of QALYs, the larger is the true growth in the health effect.

In the absence of information on the effect of NHS care on health and hence on the true growth rate it is impossible to say how large the underestimation of the overall growth rate of the effect of hospital care on health is. If our central guesstimate of the average value of $k = q^{o}_{t} / q^{*}_{t} = 0.8$ is correct, then the kind of increases in short term survival which are perhaps towards the upper end of what is plausible (from 0.970 to

0.971 or 0.972 -- corresponding to proportionate reductions in mortality of 3.3% or 6.75%) underestimate the true growth in the health effect by 0.5% to 1%. Of course, the calculation takes no account of changes in health effects arising from increases in health conditional on surviving treatment. If q^*_{jt} grows then the simple survival adjustment would be more of an underestimate. Once again we see the importance of having improved estimates of health conditional on surviving treatment.

Table 4.2 Error in using survival growth rate as estimate of growth rate in health effect of treatment

		Year t	Year t +1				
Survival		0.97	0.971	0.972	0.975	0.980	0.985
Mortality		0.03	0.029	0.028	0.025	0.020	0.015
Mortality % decrease			-3.33%	-6.67%	-16.67%	-33.33%	-50.00%
Survival % growth			0.10%	0.21%	0.52%	1.03%	1.55%
True health growth							
if true k =	0.5		0.21%	0.43%	1.06%	2.13%	3.19%
	0.8		0.59%	1.18%	2.94%	5.88%	8.82%
	0.9		1.43%	2.86%	7.14%	14.29%	21.43%
Error using survival growth							
if true k =	0.5		0.11%	0.22%	0.55%	1.10%	1.65%
	0.8		0.49%	0.97%	2.43%	4.85%	7.28%
	0.9		1.33%	2.65%	6.63%	13.25%	19.88%

If q^*_{jt} is constant ((45) holds) so that holds the difference between the quality adjusted cost weighted and survival adjusted cost weighted indices is

$$I^{xq}_{ct} - I^{xa}_{ct} = \sum_j \left(1 + g_{xjt}\right) g_{ajt} \left(\frac{q^0_{jt}}{q_{jt}}\right) \frac{x_{jt} c_{jt}}{\sum_{j'} x_{j't} c_{j't}} \qquad (46)$$

Since we do not observe q^*_{jt}, q^o_{jt} we cannot determine the magnitude of the absolute downward bias in using a_{jt+1} / a_{jt} instead of q_{jt+1} / q_{jt}. When there is efficient allocation (30) for conditions where the alternative to NHS treatment is very poor health (q^o_j is small) or treatment has a large effect so (that q^*_{jt} is large relative to q^o_j), the bias is small. But if q_{jt} is very small (the treatment has a small effect on health) the

bias is very large. Fortunately, the smaller the health gain from the treatments the smaller the weight of the treatment in the cost adjusted value weighted index (since costs are proportional to health gains by assumption (30) and so the downward bias is bounded. But if (30) does not hold so that unit cost is not proportional to marginal value then the downward bias when q_{jt} is small may not be offset by having a low cost weight attached to such outputs.

We also see from (43) that an increase in survival will have a smaller effect on the index the smaller is the cost weight c_{jt}. When (30) holds this is reasonable. The increase in the health effect from an increase in survival is proportional to q^*_{jt} and the smaller is c_{jt} the smaller is $q_{jt} = a_{jt}q^*_{jt} - q^o_{jt}$ and the more likely is q^*_{jt} to be small. But if (30) does not hold, the fact that survival gains in low cost activities will have smaller effects on the index than survival gains in high cost activities is less appealing.

A further difficulty with the pure survival adjustment is that it takes no account of the age of the patients treated. A given post operative survival gain has the same effect on the output index if the treatments have the same cost and volume, even though one treats a much younger group of patients. Again this is reasonable if unit costs are proportional to health effects since the cost weight adjusts for the effects of differences in average age at treatment on health effects. But if we do not believe that unit costs are proportional to health effects we may want to find another means of allowing for differences in age across HRGs.

We conclude that adjusting for survival is better than ignoring survival changes. We consider in the next two sections how it is possible to improve on a pure survival adjustment by combining a little more information (from a small sample of treatments where there some information on health effects and from estimates of life expectancy) with further assumptions.

4.8.2 Incorporating estimates of health effects

To proceed further we must, in the current state of information about the health effects of treatment, replace knowledge with additional assumptions. We consider the implications of assuming

(a) health conditional on treatment is constant from one period to the next for all treatments

$$q^*_{jt} = q^*_{jt+1} \tag{47}$$

(b) the ratio of health conditional on surviving treatment to health conditional on no treatment is constant over time and the same for all treatments

$$q^o_{jt} / q^*_{jt} = k \tag{48}$$

The assumption of efficient allocation (30) by itself does not enable us to claim that a simple cost weighted index is what we want:

$$I^{xq}_{yt} = \sum_j \frac{x_{jt+1} q_{jt+1}}{x_{jt} q_{jt}} \neq \sum_j \frac{x_{jt+1} c_{jt}}{x_{jt} c_{jt}} \tag{49}$$

unless the quality of care is constant. But assumption (30) is useful since we can combine it with (47) and (48) to get

$$I^{xq}_{yt} = I^{xq}_{ct} = \frac{\sum_j x_{jt+1} \left(\frac{q_{jt+1}}{q_{jt}} \right) c_{jt}}{\sum_j x_{jt} c_{jt}} = \frac{\sum_j x_{jt+1} \left(\frac{a_{jt+1} - k}{a_{jt} - k} \right) c_{jt}}{\sum_j x_{jt} c_{jt}} \tag{50}$$

Notice that we cannot make assumptions (30), (47) and (48) and then construct an index by applying the quality adjustments factors $a_{jt+1} - k$, $a_{jt} - k$ separately to the outputs in each year since

$$\frac{\sum_j x_{jt+1}[a_{jt+1} - k] c_{jt}}{\sum_j x_{jt}[a_{jt} - k] c_{jt}} = \frac{\sum_j x_{jt+1}[a_{jt+1} - k] \lambda_t \pi_t q_{jt}}{\sum_j x_{jt}[a_{jt} - k] \lambda_t \pi_t q_{jt}}$$

$$= \frac{\sum_j x_{jt+1}(q_{jt+1} / q^*_j) q_{jt}}{\sum_j x_{jt}(q_{jt} / q^*_j) q_{jt}} = \frac{\sum_j x_{jt+1} q_{jt+1} (q_{jt} / q^*_j)}{\sum_j x_{jt} q_{jt} (q_{jt} / q^*_j)} \tag{51}$$

$$\neq \frac{\sum_j x_{jt+1} q_{jt+1}}{\sum_j x_{jt} q_{jt}} = I^{xq}_{yt} \tag{52}$$

Only if q_{jt} / q^*_j is the same across all treatments would separate application of the

quality adjustments factors $a_{jt+1} - k$, $a_{jt} - k$ to x_{jt+1}, x_{jt} produce the correct result. Our assumptions imply that q_j^o / q_j^* is constant across treatments, not that q_{jt} / q_j^* is constant across treatments. The difference is that q_j^o / q_j^* does not involve the survival rate a_{jt}, whereas $q_{jt} / q_j^* = (a_{jt} q_j^* - q_j^o) / q_j^* = a_{jt} - q_j^o / q_j^* = a_{jt} - k$ does. To ensure that q_{jt} / q_j^* is equal across all treatments we would have to make the additional (and patently false) assumption that all a_{jt} are equal which then means that a_{jt+1} must be the same across all j, though possibly different from a_{jt}. Hence the separate application of the quality adjustments factors $a_{jt+1} - k$, $a_{jt} - k$ to x_{jt+1}, x_{jt} as in (51) is valid only if we can apply the same adjustment factors to all treatments which is equivalent to scaling a simple cost weighted index by $(a_{t+1} - k)/ (a_t - k)$.

From our review of the EQ5D literature and analysis of data from BUPA and York NHS Trust (summarised in Appendix C), we have snapshot estimates of health status before (h_ℓ^b) and after (h_ℓ^*) for a limited set of treatments. (See section 6 where we use these estimates in a specimen index for the set of treatments to illustrate, *inter alia*, the implications calculating a value weighted index rather than a cost weighted index.) We also use these estimates in section 5 to get very rough estimates of the cost weighted quality adjusted output index for all activities I_{ct}^{xq} by making assumptions (a) and (b) above and using h^o / h^* - the average value of h_ℓ^b / h_ℓ^* for our limited sample of treatments - as an estimate of $q_j^o / q_j^* = k$ in (50). We use k = 0.8 as our base case but consider variants 0.7 and 0.9. Notice that we are estimating a ratio of sums of discounted QALYs by a ratio of health status snapshots. We discuss the implications further in section 4.8.3.

Clearly the assumptions that the set of treatments for which we happen to have data on h_ℓ^b / h_ℓ^* are representative of the effects of all NHS treatment is very strong but we make it to illustrate the importance of having information on the health effects of treatment.

Calculating the index (50) with k set equal to the mean of the values in the sample of

procedures for which there are health outcome data creates a problem with some activities which appear to have a negative health effect given the assumed value of k and the observed value of a_{jt}. Some activities have high mortality rates so that the terms $a_{jt+1} - k$, $a_{jt} - k$ in the quality adjustment factor in (50) are close to zero or negative. Small changes in a_{jt} can then to large changes in the index and if both are negative the adjustment will indicate negative growth when there has been an improvement in output in the sense that $0 > a_{jt+1} - k > a_{jt} - k$.

A negative value of $a_{jt} - k$ implies that the activity has a negative social value. This may be true for some treatments but is clearly incorrect for others, such as terminal care. In the case of terminal care the problem arises from the factorisation of the health effects as $a_{jt} q^*_{jt} - q^o_{jt}$ where q^*_{jt} is the post treatment health stream conditional on survival. This is not appropriate for terminal care since all treatment ends in death. The solution is to reformulate the health gain as total discounted QALYs from start of treatment minus q^o_{jt}. Terminal care can then have a positive health outcome: patients are better off with terminal care than without it. In other cases the problem may be that our estimate of k_j as the mean of our sample of procedures for which there are outcome data is too large: if we had health data specific to the treatment $a_{jt} - k_j$ would be positive. Finally, it is possible that there are treatments which have a negative health effect and no other valuable characteristics: they have a negative social value. These create problems for a cost weighted index because they clearly violate the underlying assumption that their unit cost measures their social value: unit costs cannot be negative.

If we had information on health effects and could use health effect weights (as in (28)) then activities with negative or very small social value would not lead to small changes in a_{jt} having disproportionate effects on the index because their weight in the index would be negative or very small.

In the absence of such information we have to make *ad hoc* adjustments to calculate an index which is not disproportionately sensitive to changes in a_{jt} for activities with small or negative $a_{jt} - k$. We adopt a cut off rule: if either $a_{jt+1} - k$ or $a_{jt} - k$ is less than a threshold value, say 0.15, we use the pure survival adjustment a_{jt+1}/a_{jt} and

otherwise we use $(a_{jt+1} - k)/(a_{jt} - k)$. We calculate indices with various values of the cut off in section 5.4.

Table 4.3 shows the magnitude of the error in the calculated growth rate in the health effect for different size errors in the estimated value of $k = q_t^o / q_t^*$. The illustration assumes that the survival rate is 0.97 which is not far from the average post-operative survival rate. The assumed proportionate increase in survival and reduction in the mortality rate are on the large size, as is the true growth rate in the health effect when the true k exceeds 0.70. Notice again that because survival is high the survival growth rate (1.03%) is low, and is considerably less than the true growth in the health effect in the second column. The third column shows the error in adjusting purely by survival i.e. by setting $k = 0$. A pure survival adjustment does worse than assuming a positive value of k when the true value of k is 0.71 or above: it does worse when the proportionate effect of treatment is smaller. Our central estimate of k based on the small sample of HRGs where there is health effects data is 0.8 and for a true value of $k = 0.81$ we see that the pure survival adjustment is worse than setting k at 0.7, or 0.8.

Table 4.3 Effect of error in estimated $k = q_t^o / q_t^*$ on error in calculated growth rate in health effect

Year t survival 0.97, mortality 0.03; year t+1 survival 0.98, mortality 0.02; % growth in survival 1.03%, % growth in mortality -34%.

	Estimated k	0.0	0.7	0.8	0.9
	Estimated growth in health effect	1.03%	3.70%	5.88%	14.29%
True k	True growth in health effect	Estimated minus true growth rate in health effect			
0.91	16.67%	-15.64%	-12.96%	-10.78%	-2.38%
0.85	8.33%	-7.30%	-4.63%	-2.45%	5.95%
0.81	6.25%	-5.22%	-2.55%	-0.37%	8.04%
0.71	3.85%	-2.82%	-0.14%	2.04%	10.44%
0.51	2.17%	-1.14%	1.53%	3.71%	12.11%
0.31	1.52%	-0.48%	2.19%	4.37%	12.77%

The table suggests that even using the same fairly rough and ready estimate of k for

all treatments where survival exceeds *k* by a reasonably margin (say 0.05 or more) will be better than just using a pure short term survival adjustment.

4.8.3 Life expectancy and health effects

We now consider how it is possible to use information on the age of treated patients to modify the survival and health effects adjustments.

Consider a simple example. Let L_{jt} be the certain remaining length of life of patients who survive treatment *j* in year *t* and of those who are not treated. h_j^* and h_j^o are the levels of health status conditional on treatment and without treatment in all periods and these are constant over the remaining life of patients and are not affected by the period of treatment (there is no technological progress). The expected discounted health gain from treatment *j* in period *t* is

$$q_{jt} = a_{jt}\int_0^{L_{jt}} h_j^* e^{-rs} ds - \int_0^{L_{jt}} h_j^o e^{-rs} ds = \left(a_{jt}h_j^* - h_j^o\right)\left(\frac{1-e^{-rL_{jt}}}{r}\right) \qquad (53)$$

The quality adjustment factor to be applied to the between period *t* and *t* + 1 is therefore

$$\frac{q_{jt+1}}{q_{jt}} = \frac{\left(a_{jt+1}h_j^* - h_j^o\right)}{\left(a_{jt}h_j^* - h_j^o\right)}\left(\frac{1-e^{-rL_{jt+1}}}{1-e^{-rL_{jt}}}\right) \qquad (54)$$

Replacing h_j^o / h_j^* with the constant *k* which we estimate from the mean of our sample of health effect studies as in the previous section, the quality adjusted cost weighted output index analogous to (50) but allowing for changing life expectancy due to changes in the mix of patient types is

$$I_{ct}^{xa} = \frac{\sum_j x_{jt+1}c_{jt}\dfrac{\left(a_{jt+1}-k\right)}{\left(a_{jt}-k\right)}\left(\dfrac{1-e^{-rL_{jt+1}}}{1-e^{-rL_{jt}}}\right)}{\sum_j x_{jt}c_{jt}} \qquad (55)$$

Notice that in section 4.8.2 we assumed that the mean ratio of snapshot health status for our small set of specimen HRGs for which such data exists was equal to the ratio of sums of discounted QALYs over the lifetime of patients (q_j^o / q_j^*). Here we make

the possibly more plausible assumption that the ratio of snapshot health status values is equal to the ratio of snapshot health status without and with treatment (h_j^o / h_j^*).

The adjustment rests on the implicit assumption that all patients in *t* have the same life expectancy L_{jt}. Since patients generally differ by age and often by gender it will matter whether we calculate the life expectancy adjustment using estimates based on the mean age and gender of patients $e^{-rL_{jt}} = e^{-rE_i L_{ijt}}$ or whether we calculate use $e^{-rL_{ijt}}$ for each age and gender group and then use $Ee^{-rL_{ijt}}$. The difference between $Ee^{-rL_{ijt}}$ and $e^{-rL_{jt}}$ is typically very small, less than 0.5% on average for electives. In section 5 we report results using both approaches to determine how sensitive the indices are to the use of grouped or individual calculations of the life expectancy adjustment.

We use data on age specific health status from the 1996 Health Survey for England plus life tables to calculate healthy life expectancy, rather than actual life expectancy for use in the output indices. See Appendix A.

Note that if life expectancy does not change between periods the life expectancy terms in (55) cancel out. Thus the index does not reflect cross treatment differences in age at treatment. The rationale is that we have assumed that costs are proportional to the marginal value of treatment so that any differences in average age at treatment which affect life expectancy and health gains are already allowed for. Since we have suggested that this assumption is not appealing we have investigated using life expectancy with the survival adjustment and estimated health effects in our specimen index where we have HRG specific information on the health effects. We report in section 6.4 the results from estimating

$$\frac{\sum_j x_{jt+1}\left(a_{jt+1}h_j^* - h_j^o\right)(1-e^{-rL_{jt+1}})}{\sum_j x_{jt}\left(a_{jt}h_j^* - h_j^o\right)(1-e^{-rL_{jt}})} \qquad (56)$$

4.8.4 Cost of death adjustment

The conclusion from sections 4.8.1 to 4.8.3 is that in the absence of much of the

required data on the effects of NHS activity on health we have either to use a simple survival adjustment which will have a small effect or to make strong assumptions, bolstered by further ad hoc adjustments, in order to incorporate health effects into a cost weighted index. There is a third possibility to which we now turn which replaces the assumption of efficient allocation (30) with another assumption about the relationship between unit costs and the health effects of treatment.

Instead of making the assumption of efficient allocation (30) to use information on costs c_{jt} to make inferences about the effect of treatment on health suppose we assume that unit costs are equal to the value of output before making any allowance for death:

$$c_{jt} = p_t(q^*_{jt} - q^o_{jt}) = p_t(q_{jt} + m_{jt}q^*_{jt}) \tag{57}$$

This implies that health care providers take no account of mortality risk when determining treatment and only consider the gain in health from successful treatment. Then we can use the assumption to estimate the value of the true health effects, which allow for mortality risk,

$$p_t q_{jt} = p_t[(1 - m_{jt})q^*_{jt} - q^o_{jt}] = c_{jt} - m_{jt}p_t q^*_{jt} \tag{58}$$

and the value of the q_{jt+1} at period t value of health is

$$p_t q_{jt+1} = \frac{q^*_{jt+1} - q^o_{jt+1}}{q^*_{jt} - q^o_{jt}} c_{jt} - p_t m_{jt+1} q^*_{jt+1} \tag{59}$$

This gives an estimate of the value weighted index as

$$I^{\gamma}_x = \frac{\sum_j \left\{ \frac{q^*_{jt+1} - q^0_{jt+1}}{q^*_{jt} - q^0_{jt}} c_{jt} - m_{jt+1}\pi_t q^*_{jt+1} \right\} x_{jt+1}}{\sum_j \left\{ c_{jt} - m_{jt}\pi_t q^*_{jt} \right\} x_{jt}} \tag{60}$$

If nothing is known about $\frac{q^*_{jt+1} - q^0_{jt+1}}{q^*_{jt} - q^0_{jt}}$ we can set $\frac{q^*_{jt+1} - q^0_{jt+1}}{q^*_{jt} - q^0_{jt}} = 1$. This is equivalent to assuming that $q^o_{jt} = k_j q^*_{jt}$ for all $\boldsymbol{j}$ and $\boldsymbol{t}$ and $q^*_{jt} = q^*_{jt+1}$. The index then simplifies to

$$I^{\gamma}_x = \frac{\sum_j \left\{ c_{jt} - m_{jt+1}\pi_t q^*_{jt+1} \right\} x_{jt+1}}{\sum_j \left\{ c_{jt} - m_{jt}\pi_t q^*_{jt} \right\} x_{jt}} \tag{61}$$

We can interpret $p_t q^*_{jt}$ as the cost of death for a patient who would have survived treatment. Thus the index in (61) is a cost weighted output index in which we deduct a

cost of death from the unit cost of activity. We can use p_t = £30000, a value which is generally believed consistent with the approach used by NICE in order to calculate I_X^g . However, we still need an estimate of the discounted sum of QALYs obtained by the average individual who receives treatment *j* in periods *t* and *t*+1. One possibility is to assume that q_{jt}^* is proportional to the average expected discounted sum of QALYs for people with the same age and gender distribution as those treated ($\hat{q}_{jt}$). We can use general population estimates of QALYs in from the Health Survey for England, combined with appropriate life tables. However, we are still left with problem of estimating the proportionality factors $f_{jt} = q_{jt}^* / \hat{q}_{jt}$. We would expect the factor for patients receiving cataracts to differ from the factor for those undergoing heart surgery.

The only source of information from which we can infer patient health post treatment is the death rate. One apparently simple and appealing possibility is to employ a proportionality factor $f_{jt}(m_{jt}) = 1 - m_{jt}$ with the aim of ensuring that we have an index which attaches a low cost of each death to death in those treatments which have a high mortality rate. With this adjustment we obtain the index

$$I_x^\delta = \frac{\sum_j \left\{ c_{jt} - m_{jt+1}\pi_t f_{jt+1}\left(m_{jt+1}\right)\hat{q}_{jt+1} \right\} x_{jt+1}}{\sum_j \left\{ c_{jt} - m_{jt}\pi_t f_{jt}\left(m_{jt}\right)\hat{q}_{jt} \right\} x_{jt}}$$

$$= \frac{\sum_j \left\{ c_{jt} - m_{jt+1}\pi_t \left(1 - m_{jt+1}\right)\hat{q}_{jt+1} \right\} x_{jt+1}}{\sum_j \left\{ c_{jt} - m_{jt}\pi_t \left(1 - m_{jt}\right)\hat{q}_{jt} \right\} x_{jt}} \qquad (62)$$

where $\hat{q}_{jt}$ is estimated from life tables, HSE QALY estimates and the age-gender distribution of patients getting treatment *j*. Since very few mortality rates exceed 0.5 we can avoid the difficulty that the cost of death is decreasing with the mortality rate when it exceeds 0.5 by setting $f_{jt} = 0$ when $m_{jt} > 0.5$.

We experimented with calculations of the index based on a value of £30,000 per QALY. We found that the simple proportionality factor f_{jt} = 1- m_{jt} point to the hospital service as a whole subtracting output. There is no practical resolution to this bizarre result except to use a scaling factor which several orders of magnitude smaller than 1-*m*. This would be entirely arbitrary. In addition, the underlying assumption

(57) about the relationship between unit costs and health effects on which the cost of death adjustment rests is less acceptable than the efficient allocation assumption made elsewhere in our this report. We do not recommend this approach.

4.8.5 In-hospital versus 30 day mortality

Hospital Episode Statistics have a field indicating whether the patient was dead or alive on discharge from hospital. The data are available from 1988/89 onwards. It is also possible, though it requires considerable processing, to match HES records with ONS mortality records to count mortality within any required period after admission to hospital. HES has recently introduced a field recording the date of death if the date was between the start of the HES year (1 April) and 30 April of the following year (30 days after the end of the HES year). Thus is it possible to count deaths in hospital plus those within 30 days of discharge. We assign deaths to HRGs using the HRG of the first episode of the spell where a spell consists of more than one episode.

There are three obvious counts of deaths: in hospital deaths, deaths within 30 days of admission, in hospital deaths plus deaths within 30 days of discharge. Measuring deaths within 30 days of discharge will represent a longer follow up than deaths within 30 days of admission.

Some deaths (e.g. road accidents) outside hospital will have nothing to do with the quality of NHS care. Moreover, the matching of HES to ONS is not perfect: around 10% of spells with a discharged dead code according to HES do not have a matching ONS death record (see Appendix B).

This suggests that some patients discharged alive but dying within 30 days may not have a matching ONS death record. Hence using in hospital deaths from HES plus deaths within 30 days after discharge from ONS will understate the 30 day post discharge mortality rate. On the other hand counting only deaths in hospital runs the risk of missing deaths which occur outside hospital which are capable of being affected by the quality of care.

The correlations between in-hospital and 30 day post discharge survival rates for all HRGs in 2002/03 are 0.985 for electives and 0.991 for non-electives. (The survival rates are based on CIPS since this is our preferred unit of output for the NHS hospital sector.) The correlations between the growth rates of the two measures of survival rates (in-hospital and 30 day), taken between 2001/02 and 2002/03 are 0.944 for electives and 0.994 for non-electives.

Figures 4.2 and 4.3 plot the growth rates of in-hospital and 30-day survival rates between 2001/02 and 2002/03 for elective HRGs and for non-elective HRGs. These growth rates have been Winsorised at the 5th and 95th percentile to remove outlier observations.

Figure 4.2 Growth rates for in-hospital and 30 day CIPS based survival rates, 2001/02-2002/03, electives

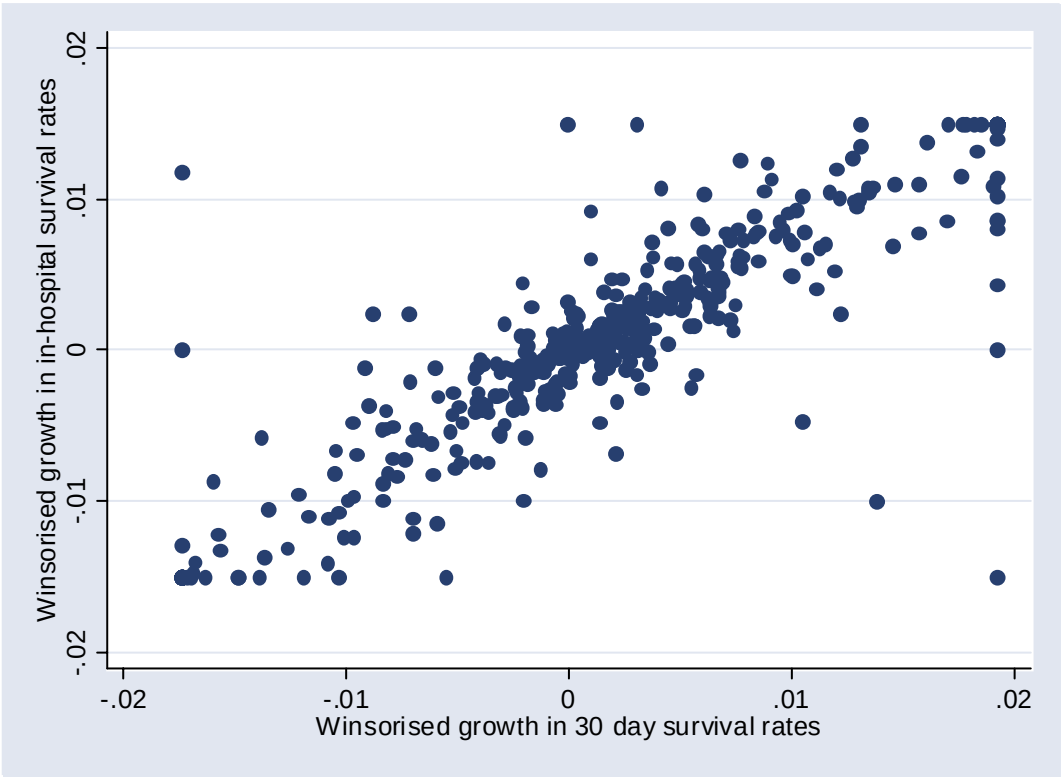
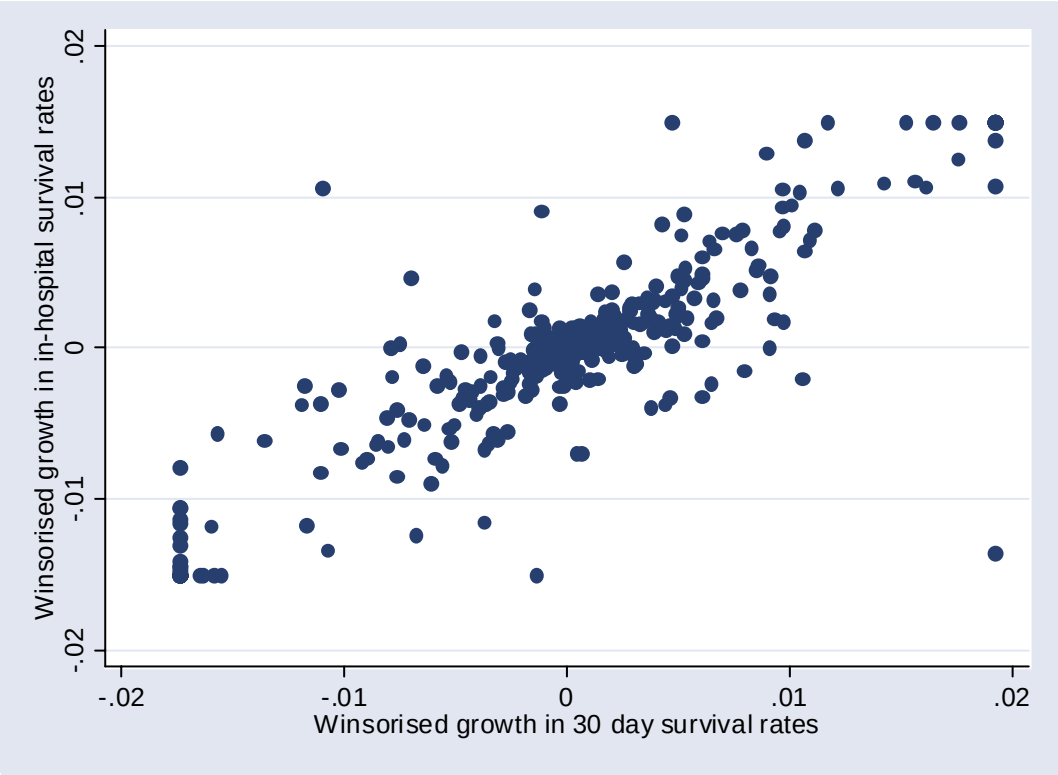


Figure 4.3 Growth rates for in-hospital and 30 day survival rates, 2001/02-2002/03, non-electives



Some HRGs have small numbers of cases (see Appendix B) so that their death rates are subject to large random fluctuations. It would be possible to allow for small number randomness with various shrinkage estimators but we decided not to do so. Precisely because such HRGs have small amounts of activity they will have little influence on the survival adjusted indices as they will account for a tiny proportion of activity.

We have estimated indices with both types of death rate in section 5. We feel that at the moment the choice between possibly more accurately recorded in hospital deaths and the possibly more useful but less well measured 30 day deaths is finely balanced but have a mild preference for 30 day mortality. The data on 30 day deaths should continue to improve. Moreover, there is some evidence that publication of in hospital mortality rates in US led to reductions in reported in-hospital mortality for some conditions but an increases in reported 30 day deaths (Baker *et al*., 2002). Use of mortality data to quality adjust an output series does not necessitate its use as a performance indicator but it seems more prudent to use a mortality measure which is

less susceptible to manipulation.

We recommend that the DH continue to encourage the refinement of the record linkage and that the short term survival adjustment be based on 30 day deaths.

4.8.6 Conclusions: survival based quality adjustment

In this sub section we have

- constructed a set of quality adjustments to the cost weighted output index which attempt to allow for the changing health effects of treatment.
- shown how to use estimates on long term survival (say up to five years) when these become available (section 4.7).
- shown how to use existing measures of short term survival (section 4.8.1).
- demonstrated that the pure survival adjustment will almost certainly underestimate the true growth in the effect of treatment on health,
- shown how guesstimates of the proportional effect of treatment compared to no treatment (section 4.8.2) and data on life expectancy (section 4.8.3) can be incorporated into survival adjustment.
- shown that assuming that the unit costs are proportional to health effects when mortality risk is ignored, rather than making the standard assumption that allocation is efficient and unit costs proportional to health leads to an adjustment which takes the form of a deduction of a cost of death from the output valued using the unit costs (section 4.8.4).

We defer judgement on recommending one of these adjustments until we have considered the results from calculating them on actual data, either on the whole of the hospital sector (section 5) or for our specimen set of HRGs for which we have some health effect data (section 6). Section 7 contains calculations based on our preferred variant. We wish to see if the resulting estimates of output growth are either implausible or overly sensitive to unverifiable assumptions about the parameters. However in light of the dubious underlying assumptions about unit costs required to derive the cost of death adjustment (section 4.8.4) and some preliminary calculations we do not recommend it and so do not report results using it.

Table 4.4 Summary of output indices with survival based adjustments

Quality adjustments	Weights	Form	Rationale in section	Results in section	Assumptions	Comments
Long term survival	Costs	$\dfrac{\sum_j x_{jt+1}c_{jt}\dfrac{a_{jt+1}}{a_{jt}}\sum_{s=1}^{S}\left(\dfrac{\sigma^*_{jt+1}(s)}{\sigma^*_{jt}(s)}\right)\left(\dfrac{\sigma^*_{jt}\delta^s}{\sum_{s=1}^{5}\sigma^*_{jt}(s)\delta^s}\right)}{\sum_j x_{jt}c_{jt}}$	4.7	Not yet feasible	Efficient allocation. $h^*_{jt}(s)$ constant s = 1,…,5; zero s >5. $h^o_{jt}(s)$ = 0 all s.	Survival increases in more costly HRGs have bigger impact.
Survival.	Costs	$\dfrac{\sum_j c_{jt}x_{jt+1}\left(\dfrac{a_{jt+1}}{a_{jt}}\right)}{\sum_j c_{jt}x_{jt}}$	4.8.1	5.4.1, 6.3	Efficient allocation.	Survival growth has larger effect if HRG more costly. Underestimates true growth. Impact of survival unaffected by age of those treated.
Survival. Health effect.	Costs	$\dfrac{\sum_j x_{jt+1}\left(\dfrac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)c_{jt}}{\sum_j x_{jt}c_{jt}}$	4.8.2	5.4.2, 6.4	Efficient allocation. q^*_{jt}, q^o_{jt} constant. Same proportionate effect of treatment on QALYs with and without treatment, all HRGs ($q^o_j / q^*_j = k$) in sec 5, 6. k_j varies across j in section, 7.	Quality growth has larger effect if HRG more costly. Quality adjustment unaffected by age of those treated. Requires adjustment set to a_{jt+1}/a_{jt} survival rate low (close to k) to avoid instability in index
Survival. Health effect. Life expectancy.	Costs	$\dfrac{\sum_j x_{jt+1}c_{jt}\dfrac{\left(a_{jt+1}-k\right)}{\left(a_{jt}-k\right)}\left(\dfrac{1-e^{-rL_{jt+1}}}{1-e^{-rL_{jt}}}\right)}{\sum_j x_{jt}c_{jt}}$	4.8.3	5.4.3	Efficient allocation. h^*_{jt}, h^o_{jt} constant; same proportionate effect of treatment on health status all HRGs in sec 5, 6. k_j varies across j in section, 7. Certain life; same for treated and untreated; same for all patients in HRG	Quality growth has larger effect if HRG more costly. Requires adjustment set to a_{jt+1}/a_{jt} survival rate low (close to k) to avoid instability in index

Survival. Life expectancy.	Life expectancy	$\frac{\sum_j x_{jt+1}\left(a_{jt+1}h_j^* - h_j^o\right)\left(1-e^{-rL_{jt+1}}\right)}{\sum_j x_{jt}\left(a_{jt}h_j^* - h_j^o\right)\left(1-e^{-rL_{jt}}\right)}$	4.8.3	6.4	h_{jt}^*, h_{jt}^o constant. Certain life; same for treated and untreated; same for all patients in HRG	
Cost of death	Costs	$\frac{\sum_j \left\{c_{jt} - m_{jt+1}\pi_t f_{jt+1}\left(m_{jt+1}\right)\hat{q}_{jt+1}\right\}x_{jt+1}}{\sum_j \left\{c_{jt} - m_{jt}\pi_t f_{jt}\left(m_{jt}\right)\hat{q}_{jt}\right\}x_{jt}}$	4.8.5	Not reported	Cost proportional to health effect ignoring mortality risk	Requires proportionality factor f_{jt} to be very small to avoid negative output

c_{jt} unit cost, volume x_{jt} volume; a_{jt}, m_{lt} proportion patients alive, dead on discharge (or after 30 days); h_j^* constant health status conditional on surviving treatment. h_j^o constant health status if not treated. $\sigma_t^*(s)$ probability of surviving s years; k estimate of proportionate effect of treatment on quality adjusted life years (QALYs); L_{jt} life expectancy at mean age of patients treated; r discount rate on QALYs; π_t value of QALY (£s); $\hat{q}_{jt}$ QALYs lost by death of average patient; $f_{jt}(m_{jt})$ proportionality factor measuring seriousness of HRG.

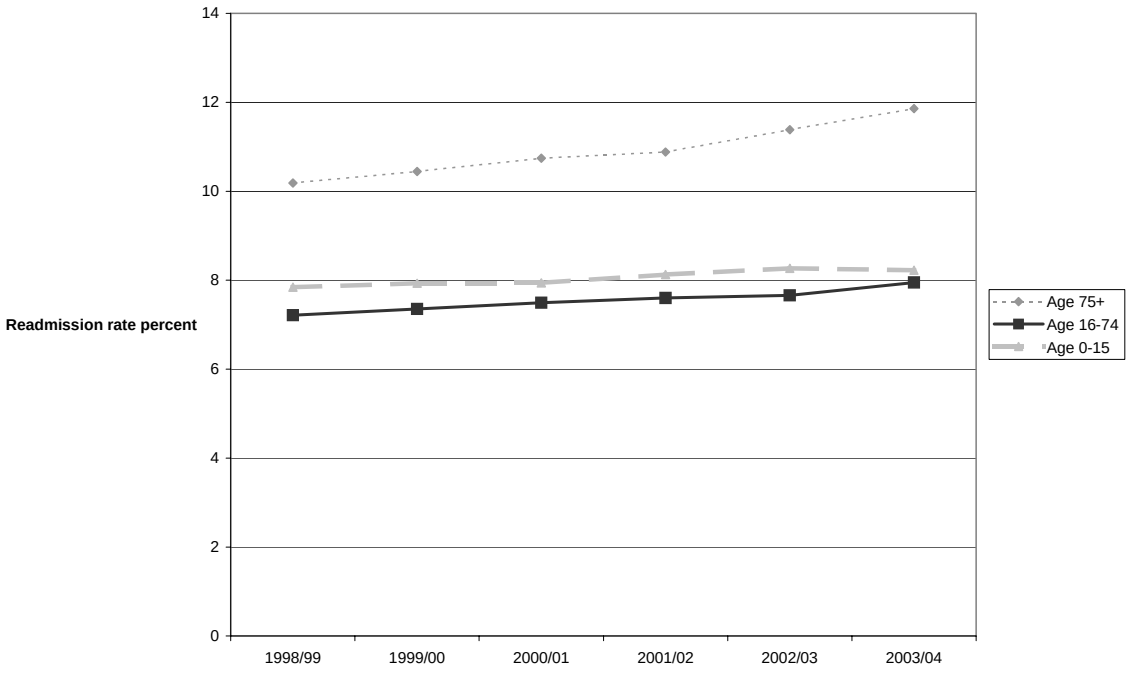
4.9 Readmissions

4.9.1 Readmissions as health effects

A non-trivial proportion of patients are readmitted to hospital within 28 days of discharge as emergencies. Figure 4.4 shows a rise in these readmissions over time, with the data being described in more detail in the Appendix. Some of these readmissions will reflect poor quality care, in that with proper care the readmission would not have occurred (Hofer and Hayward, 1995; Ludke *et al*, 1993; Thomas, 1996). However, it is not possible from these data to distinguish whether some failure of treatment occurred during the first hospital stay because "readmissions" are defined as any emergency admission at all within 28 days of discharge, whether or not they were related to the earlier admission.

Since 2001 readmission rates have been used by the Department of Health, CHI and the Healthcare Commission, as performance indicators for NHS Trusts. In this section we discuss whether and how readmissions should be used to quality adjust a cost weighted output index to reflect poor quality care on first admission. Since readmissions may also be a distressing experience for patients per se, irrespective of the consequences of poor care at first admission for their health, we also consider in section 4.9.2 how to adjust for readmissions as an aspect of the patient experience.

Figure 4.4 Emergency readmission to hospital within 28 days of discharge, by age band



Source: Department of Health

For some procedures a proportion of patients will be readmitted because of complications arising from the treatment. Let x_{jt}^1 be the number of first admissions and x_{jt}^2 be the number of readmission, so that the readmission rate is $R_{jt} = x_{jt}^2 / x_{jt}^1$. The readmissions may be in a different HRG but this is allowed for in the notation since x_{jt}^2 is the number of readmissions to whatever is the appropriate HRG for readmissions for unsuccessful treatments from the initial HRG j.

If the first admission is successful the health gain to the patient is $q_{jt}^{1*} - q_{jt}^{1o}$ (assume for the moment that there is no mortality risk for admissions or readmissions). If the operation is unsuccessful (probability F_{jt}) and they are not readmitted the effect of the first admission for this unsuccessful group $q_{jt}^{2o} - q_{jt}^{1o}$. But if they are readmitted they get a health gain from readmission, compared to having no readmission, of $q_{jt}^{2*} - q_{jt}^{2o}$. We assume that the interval between admission and readmission is short enough to ignore any effect of an unsuccessful first admission on health for the time between admissions or that health change after an unsuccessful first admission is zero.

The value of total health gain is

$$\sum_j \pi_t \left[x_{jt}^1 (q_{jt}^{1*} - q_{jt}^{1o})(1 - F_{jt}) + x_{jt}^1 F_{jt} (q_{jt}^{2o} - q_{jt}^{1o}) \; + x_{jt}^2 (q_{jt}^{2*} - q_{jt}^{2o}) \right] \qquad (63)$$

If the NHS maximises the value of health subject to the constraints that total cost does not exceed the NHS budget and that $F_{jt} x_{jt}^1 - x_{jt}^2 \geq 0$, then

$$c_{jt}^1 = \lambda_t \pi_t \left[\left(1 - F_{jt}\right)\left(q_{jt}^{1*} - q_{jt}^{1o}\right) + F_{jt}\left(q_{jt}^{2o} - q_{jt}^{1o}\right) + F_{jt}\mu_{jt} \right] \qquad (64)$$

$$\begin{aligned} c_{jt}^2 &= \lambda_t \pi_t \left[\left(q_{jt}^{2*} - q_{jt}^{2o}\right) - \mu_{jt} \right] \\ &= \lambda_t \pi_t \left[\left(1 - R_t\right)\left(q_{jt}^{1*} - q_{jt}^{1o}\right) + R_t\left(q_{jt}^{2*} - q_{jt}^{1o}\right) \right] - R_t c_{jt}^2 \end{aligned} \qquad (65)$$

where μ_{jt} is the Lagrange multiplier on $F_{jt} x_{jt}^1 - x_{jt}^2 \geq 0$ and λ_t is the reciprocal of the multiplier on the budget constraint. Consider solutions in which all those for whom the first treatment is not successful are readmitted $F_{jt} = R_{jt}$ so that the constraint binds and $\mu_{jt} > 0$. (This makes no essential difference to the conclusions.) We can substitute for μ_{jt} in (64) to get

$$c_{jt}^1 = \lambda_t \pi_t \left[\left(1 - R_t\right)\left(q_{jt}^{1*} - q_{jt}^{1o}\right) + R_t\left(q_{jt}^{2*} - q_{jt}^{1o}\right) \right] - R_t c_{jt}^2 \qquad (66)$$

In allocating resources to first admissions allowance must be made for the fact that with probability $F_{jt} = R_{jt}$ the patient will have an unsuccessful first admission and then be readmitted, getting a health gain $q_{jt}^{2*} - q_{jt}^{1o}$ but generating additional costs c_{jt}^2. The health gain when $F_{jt} = R_{jt}$ is

$$\sum_j \pi_t \left[x_{jt}^1 (q_{jt}^{1*} - q_{jt}^{1o})(1 - R_{jt}) + x_{jt}^2 (q_{jt}^{2*} - q_{jt}^{1o}) \right] \qquad (67)$$

Current practice is to calculate a cost weighted outcome index (CWOI) which includes both first and subsequent admissions

$$\frac{\sum_j \left[x_{jt+1}^1 c_{jt}^1 + x_{jt+1}^2 c_{jt}^2 \right]}{\sum_J \left[x_{jt}^1 c_{jt}^1 + x_{jt}^2 c_{jt}^2 \right]} \qquad (68)$$

If there is efficient allocation ((65) and (66) hold) and if there is no change in health effects or in the readmission rate then the CWOI (68) equals the value weighted output index

$$\frac{\sum_j \left[x^1_{jt+1} c^1_{jt} + x^2_{jt+1} c^2_{jt} \right]}{\sum_J \left[x^1_{jt} c^1_{jt} + x^2_{jt} c^2_{jt} \right]} = \frac{\sum_j \pi_t \left[x^1_{jt+1} (q^{1*}_{jt+1} - q^{1o}_{jt+1})(1 - R_{jt+1}) + x^2_{jt+1} (q^{2*}_{jt+1} - q^{1o}_{jt+1}) \right]}{\sum_j \pi_t \left[x^1_{jt} (q^{1*}_{jt} - q^{1o}_{jt})(1 - R_{jt}) + x^2_{jt} (q^{2*}_{jt} - q^{1o}_{jt}) \right]} \quad (69)$$

To show this use (65) and (66) and write the denominator in (68) is

$$\sum_j \lambda_t \pi_t \left\{ x^1_{jt} \left[\left(1 - R_{jt}\right)\left(q^{1*}_{jt} - q^{1o}_{jt}\right) + R_{jt}\mu_{jt} \right] + x^2_{jt} \left[\left(q^{2*}_{jt} - q^{1o}_{jt}\right) - \mu_{jt} \right] \right\}$$

$$= \sum_j \lambda_t \pi_t \left\{ x^1_{jt} \left[\left(1 - R_{jt}\right)\left(q^{1*}_{jt} - q^{1o}_{jt}\right) \right] + x^2_{jt} \left(q^{2*}_{jt} - q^{1o}_{jt}\right) \right\} \quad (70)$$

which is proportional to the value of period *t* health output at the period *t* price of health. Each term in the denominator of (68) can be written as

$$x^1_{jt+1} \left(c^1_{jt} + R_{jt+1} c^2_{jt}\right) = x^1_{jt+1} \lambda_t \pi_t \left\{ \left[\left(1 - R_{jt}\right)\left(q^{1*}_{jt} - q^{1o}_{jt}\right) + R_{jt}\mu_{jt} \right] + R_{jt+1} \left[\left(q^{2*}_{jt} - q^{1o}_{jt}\right) - \mu_{jt} \right] \right\}$$

which, with the assumption of a constant readmission rate and constant health effects, is

$$\lambda_t \pi_t \left[x^1_{jt+1} \left(1 - R_{jt+1}\right)\left(q^{1*}_{jt+1} - q^{1o}_{jt+1}\right) + x^2_{jt+1} \left(q^{2*}_{jt+1} - q^{1o}_{jt+1}\right) \right] \quad (71)$$

Hence the numerator in (68) is λ_t time the value of health output in period *t* +1 at period *t* price of health.

Thus no adjustment need be made to the cost weighted output index for readmissions if there is efficient allocation and if there are no change in health effects or readmission rates. This is just a particular example of the result we noted in section 2.7. The conclusion holds if we relax the assumption that the time period between first and second admission is very short so that the change in health over this time interval can be ignored. It also holds if we allow for a non-zero mortality risk. What matters is the assumption, underlying the use of the cost weighted output index, that allocation of resources is efficient.

Notice that the conclusion does not mean that readmissions do not affect the quality of care, nor does it mean that performance indicators for Trusts based on readmission rates do measure anything of interest. Increases in readmission rates, ceteris paribus, reduce welfare. But on the assumptions made above readmissions are already allowed for in the simple cost weighted index.

If health effects change but readmission rates do not, then the quality adjustment of the index raises exactly the same issues as discussed in earlier sections. Suppose that

readmission rates change but that the health effects $q_{jt}^{1*} - q_{jt}^{1o}$, $q_{jt}^{2*} - q_{jt}^{1o}$ do not, and that the assumption of efficient allocation holds. Is there an adjustment to the CWOI which ensures that it also measures the change in the value of health effects?

We require adjustment factors ϕ which ensure that

$$\phi_1(\cdot)c_{jt}^1 + \phi_2(\cdot)R_{jt+1}c_{jt}^2 = \lambda_t \pi_t \left[\left(q_{jt+1}^{1*} - q_{jt+1}^{1o} \right) \left(1 - R_{jt+1} \right) + \left(q_{jt+1}^{2*} - q_{jt+1}^{1o} \right) R_{jt+1} \right] \quad (72)$$

where the adjustment factors can depend only on variables we know (or assume). In general the adjustment factors will depend on the health effects and, if we are unwilling to make any assumptions about them, there is no set of adjustment factors depending on cost and readmission rates which will satisfy (72). But if we are willing to make assumptions about $q_{jt}^{1*} - q_{jt}^{1o}$, $q_{jt}^{2*} - q_{jt}^{1o}$ then we can make a little progress. Suppose that:

$$\left(q_{jt}^{2*} - q_{jt}^{1o} \right) = \hat{k}_j \left(q_{jt}^{1*} - q_{jt}^{1o} \right) \qquad \left(q_{jt+1}^{2*} - q_{jt+1}^{1o} \right) = \hat{k}_j \left(q_{jt+1}^{1*} - q_{jt+1}^{1o} \right) \quad (73)$$

then by substituting (65) and (66) in the left hand side of (72) we can show that the required adjustment factors are

$$\phi_1 = \frac{(1 - R_{jt+1}) + \hat{k}_j R_{jt+1}}{(1 - R_{jt}) + \hat{k}_j R_{jt}}, \quad \phi_2 = \phi_1 \frac{R_{jt}}{R_{jt+1}} \quad (74)$$

Thus for example if $\hat{k}_j = 1$ so that readmission treatment has the same health effect as the successful first treatment

$$\frac{\sum_j \left[x_{jt+1}^1 c_{jt}^1 + x_{jt+1}^2 \frac{R_{jt}}{R_{jt+1}} c_{jt}^2 \right]}{\sum_J \left[x_{jt}^1 c_{jt}^1 + x_{jt}^2 c_{jt}^2 \right]}$$

$$= \frac{\sum_j \left[x_{jt+1}^1 \left(1 - R_{jt+1} \right) \left(q_{jt+1}^{1*} - q_{jt+1}^{1o} \right) + x_{jt+1}^2 \left(q_{jt+1}^{2*} - q_{jt+1}^{1o} \right) \right]}{\sum_J \left[x_{jt}^1 \left(1 - R_{jt} \right) \left(q_{jt}^{1*} - q_{jt}^{1o} \right) + x_{jt}^2 \left(q_{jt}^{2*} - q_{jt}^{1o} \right) \right]} \quad (75)$$

and the CWOI with readmission costs scaled by R_{jt}/R_{jt+1} is identical to the value weighted health effect index. The intuition is that if the health gain from readmission is the same as the health gain from a successful first admission then the only effect of a change in the readmission rate is via the increase in the total cost of readmissions. If $R_{jt+1} < R_{jt}$ then output, other things equal, must have increased: the cost saving on readmissions can be used to produce more health from a given budget.

If we were to take account of readmission rates in our comprehensive index, we would require readmission rates R_{jt} for HRG j in period t. This data is not routinely generated and at present is not readily usable in a productivity index in the way required for the adjustment in (75)

First, since HRGs are assigned at episode level and having linked episodes into spells (CIPS), a spell may have more than one HRG within it. A readmission may or may not be related to any of the HRGs which occur in the index admission. There is as yet no agreed method for assigning an HRG to a CIPS where there are possibly multiple episodes and multiple diagnoses. In the construction of CIPS in this report, we have assigned the HRG to the first episode of care in the admission, but as mentioned, a readmission may not be related to the first HRG, or any others in the CIP spell. A spell-based HRG assignment methodology has been developed for Payment by Results which looks for the 'dominant' procedure / diagnosis / HRG across episodes, but this is based on a provider spell rather than CIPS and again, the readmission may not be linked to the 'dominant' HRG.

Second, it is impossible, with available data, to distinguish between readmissions due to poor treatment (the element we wish to capture in the index) and readmissions due to patients being generally sick and prone to getting ill. One way of trying to make a judgement on this may be to examine whether a readmission is linked to a previous discharge, but this will be extremely complex and would need to be considered for each HRG. Indeed, the readmission could be related to any of the conditions in the spell, not necessarily the HRG coded in the CIP spell. It is possible for a readmission to be due to neglect or poor care and yet apparently entirely unrelated to the set of HRGs in the index admission.

We conclude that the data do not yet support the kind of adjustment to reflect the health effects of treatment considered in this section. A more fundamental objection may be that the method rests on the standard assumption that unit costs are proportional to the value of treatments. In the case of readmissions this assumption is more than usually questionable. In the next section we sketch a more ad hoc, less theoretically grounded method which might command more support, though it also subject to the same data problems.

4.9.2 Readmissions (and clinical errors and MRSA) as a deadweight loss

An alternative view of readmissions, which can also be applied to clinical errors and MRSA, can in principle be constructed within the confines of the cost weighted index, if the extra costs associated with these can be identified. The general principle of the cost weighted index is that costs are proportional to benefits. If follows from this that if there are some components of cost which are unrelated to benefits, then these should be omitted from the index. In other words if the costs of treating MRSA, directly associated with readmission as a consequence of plainly premature discharge or clinical errors could be identified these should be omitted from the quality-augmented cost weighted index.

The HRG costs include those elements of cost which are unrelated to benefit but which instead arise from poor provision of medical services. This means that in order to calculate the change in the value of output at constant prices we should deduct from both the numerator and denominator the costs which represent money wasted. If one wanted to represent not only the money wasted but also the disutility arising from such activity, then these deductions should be augmented.

In this exercise there is a risk of double-counting. If either MRSA or premature discharge affects mortality rates, then this is already taken into account. In the quality adjusted index. Thus the index set out here can be defended only if one is sure that this is not the case.

If we take an output, we then deduct from the value of output in constant prices the expenditures on bads arising from poor provision. We assume that bad j has a cost c^b_{jt} associated with it, and that there are x^b_{jt} examples of bad j. The index then becomes, in the example of the survival and life expectancy adjusted index

$$\frac{\sum_j x_{jt+1}c_{jt}\dfrac{\left(a_{jt+1}-k\right)}{\left(a_{jt}-k\right)}\left(\dfrac{1-e^{-rL_{jt+1}}}{1-e^{-rL_{jt}}}\right)-\sum_j x^b_{jt+1}c^b_{jt}}{\sum_j x_{jt}c_{jt}-\sum_j x^b_{jt}c^b_{jt}} \qquad (76)$$

There is an obvious similarity with the cost of death index of section 4.8.4, but with

the important difference that the cost of death has to be inferred as best one can from information about the value put on life, while the costs associated with poor service provision can in principle be identified through cost accounting.

We do not recommend this adjustment given the poor state of the data on readmissions, and the costs of associated with readmissions and MRSA. We do illustrate the effect of this type of adjustment in section 5 using data on readmissions and MRSA. We could find no usable data on clinical errors but in principle they could be incorporated in the same way.

4.10 Waiting times

Waits for diagnostic tests and treatment may affect individuals in two ways. First, they may dislike waiting per se irrespective of the effect of treatment on discounted sum of their quality adjusted life year (q_{jt}). Thus waiting time is regarded as a separate *characteristic* of health care, distinct from its effect on health. Second, longer waits can reduce the health gain from treatment and the waiting adjustment is akin to a *scaling factor* multiplying the health effect.

Although currently the data do not exist to provide satisfactory estimates of the first type of effect of waiting, section 4.10.1 sets out how such data can be used when it becomes available. We then turn in subsequent sections to consider how estimates of the second type of effect can be made with current data.

The total wait is the sum of the wait for a first outpatient appointment after referral from a GP, plus possible further waits for subsequent outpatient appointments for results of tests, plus the wait from the date the patient is placed on the waiting list for inpatient treatment. We ignore these distinct components of the total waiting time because currently data do not permit tracking of individual patients in order to calculate their total time waited. This limitation is likely to be rectified in future with data being collected in order to monitor achievement of the 18 week target for total waiting times. But given current limitations, in our empirical analysis we assess waiting time after a patient has been placed on the list for an inpatient admission. In

section 4.10.5 how to include an outpatient waiting time adjustment.

4.10.1 Waiting time as a characteristic

The first way to quality adjust the output index to reflect changing waiting times is to use direct monetary valuation of reductions in waiting times. Let π_{wjt} be the value of a reduction of one day in waiting time for treatment j in year t. The value of the output of j in year t is the sum of the values of its characteristics: health gain (q_{hjt}) and waiting time ($q_{wjt} = w_{jt}$) :

$$y_{jt} = x_{jt} p_{jt} = x_{jt}[\pi_{ht} q_{hjt} + \hat{\pi}_{wjt} q_{wjt}] = x_{jt}[\pi_{ht} q_{hjt} - \pi_{wjt} w_{jt}] \tag{77}$$

Notice that we assume that the value of a day's wait depends on the treatment waited for so that (77) differs from (10). The value weighted index (12) becomes

$$\frac{\sum_j x_{jt+1}\left[\pi_{ht} q_{hjt+1} - \pi_{wjt} w_{jt+1}\right]}{\sum_j x_{jt}\left[\pi_{ht} q_{hjt} - \pi_{wjt} w_{jt}\right]} \tag{78}$$

We report in section 6 calculations of (78) based on estimates of the health effect for a small subset of HRGs in a specimen index.

If we make the assumption of efficient allocation we can derive the *quality adjusted cost weighted index* where we take account of both health and waiting time as characteristics is a special case of (21)

$$\begin{aligned} I_{ct}^{xq} &= \sum_j \frac{x_{jt+1}}{x_{jt}} \frac{\left[\pi_{ht} q_{jt+1} - \pi_{wjt} w_{jt+1}\right]}{\left[\pi_{ht} q_{jt} - \pi_{wjt} w_{jt}\right]} \frac{c_{jt} x_{jt}}{\sum_j c_{jt} x_{jt}} \\ &= \sum_j \left\{ \frac{x_{jt+1}}{x_{jt}} \frac{q_{jt+1}}{q_{jt}} \frac{\pi_{ht} q_{jt}}{\left[\pi_{ht} q_{jt} - \pi_{wjt} w_{jt}\right]} - \frac{x_{jt+1}}{x_{jt}} \frac{w_{jt+1}}{w_{jt}} \frac{\pi_{wjt} w_{jt}}{\left[\pi_{ht} q_{jt} - \pi_{wjt} w_{jt}\right]} \right\} \frac{c_{jt} x_{jt}}{\sum_j c_{jt} x_{jt}} \\ &= \sum_j \left(1 + g_{xjt}\right) \left[(1 + g_{qhjt}) \frac{\pi_{ht} q_{hjt}}{p_{jt}} - (1 + g_{qwjt}) \frac{\pi_{wjt} w_{jt}}{p_{jt}} \right] \omega_{ct}^{jt} \end{aligned} \tag{79}$$

This index requires data on the amount of health gain q_{jt} for each treatment in order to calculate the value share $\pi_{ht} q_{hjt} / p_{jt}$ due to health. We do not have such data for all HRGs and the survival adjustments considered in section 4.8 are inadequate since they used to estimate growth rates not levels of health. Thus with available data we

cannot calculate (79). The kind of data on health before and after treatment that we recommend in section 10 would enable such an index to be calculated.

We have calculated a specimen index for (79) for the small subset of treatments where we have snapshot estimates of the health effects (see section 6). We require the monetary value of health (π_{ht}) and of waiting times (π_{wjt}) to calculate this index. For π_{ht} we use the monetary value of the QALY implied by the decisions of public bodies, such as the National Institute of Clinical Excellence and the Department for Transport (see Devlin and Parkin, 2004; Carthy *et al*., 1999). We estimate the value of waiting time from the studies reported in Ryan *et al.* (2004). Note that, even with π_{wt} of the order of £10 per day, since waiting times are short compared with the horizon over which health gains are enjoyed, the effect of the waiting time adjustment may not be large.

It is possible to calculate (79) without data on the magnitude of the health effects if we are willing to make further assumptions about the relative importance of health and waiting time characteristics. Thus if we believe that the health effect of treatment is say 10 times as important as the waiting time then we can set $\pi_{ht} q_{hjt} / p_{jt} = 10/9$ which implies $\left|\pi_{wt} w_{jt} / p_{jt}\right| = 1/9$ since the sum of the value shares must be 1. We have not done so in this report because we can think of no sensible method of estimating the relative importance of health and waiting times characteristics in the absence of data on health effects and the relative marginal social values of health and waiting times.

4.10.2 Waiting time as a scaling factor

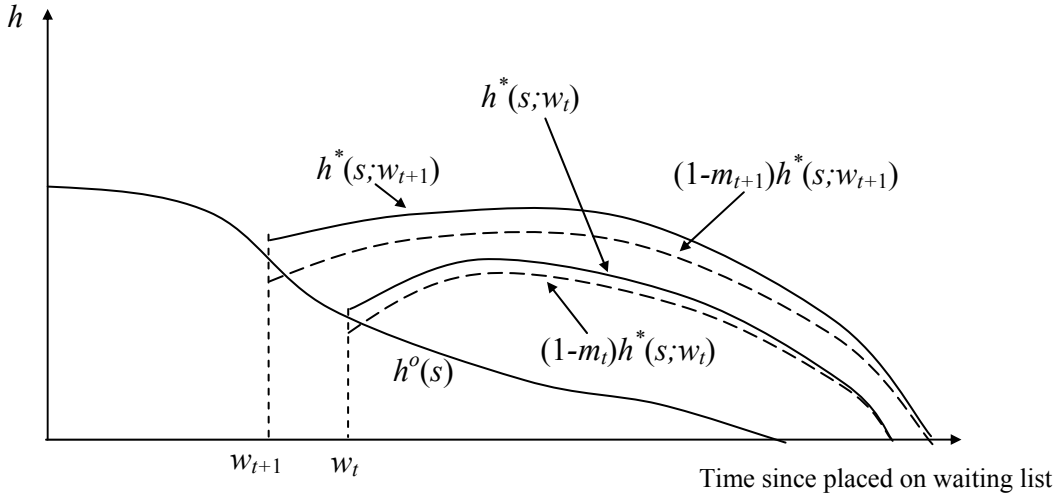
Delay may also lead to a reduced health effect q_{jt} from treatment. This can arise because (a) the condition of the patient deteriorates whilst they wait; (b) the post treatment level of health status may be reduced; and (c) if treatment is delayed the time over which the benefits from treatment accrue may be reduced.

Figure 4.5 illustrates the effect of reduction in waiting times on the health effects of

treatment. Initially we drop the subscripts for characteristics (since only health gain matters) and for type of treatment. $h^{o}(s)$ is the without treatment health status at time s (time is measured from the date at which the patient begins to wait). The waiting time in year t is w_t and the time path of health with the treatment is $h^{*}(s;w_t)$. In the following year the waiting time falls to w_{t+1} and the health time path is $h^{*}(s;w_{t+1})$. We assume that the change in the time path is due solely to the reduced wait for treatment, not to changes in technology. The dashed lines show the expected time streams given the treatment mortality probability and allow for a possible effect of reduced wait on the mortality probability. In Figure 4.5 the fact that $h^{o}(s)$ is declining means that earlier treatment prevents a deterioration in health. The fact that $h^{*}(s;w_{t+1}) > h^{*}(s;w_t)$ implies that earlier treatment leads to better post treatment health. The fact that the individual lives for longer post treatment means that they have longer to enjoy their improved health.

A recent survey of the literature (Hurst and Siciliani, 2003) found some evidence on deterioration and premature death associated with waiting for cardiology treatment but little for other procedures. Clinical reassessment of patients on a waiting list was thought to contribute to reduction in adverse outcomes of waiting but there is little data on the frequency or efficiency of re-classification of patients on waiting lists. If the NHS begins the routine collection of data on health related quality of life, the QALY improvement due to reduced waiting time should be captured by trend changes in QALYs.

Figure 4.5 Effect of reduced waiting times on time streams of health



There are two possible ways to measure the effect of treatment and hence of the effect of changes in waiting times. The first is to value treatment at the time the patient is placed on the waiting list: health effects are therefore discounted to this date. The second is to value treatment at the date it is actually received: health effects are discounted to the date of treatment. If we do not wish to quality adjust with respect to waiting times or if we are dealing with non-electives we do not need to take account of the difference between the two approaches.

4.10.2.1 Discounting to start of wait

If the patient is placed on the waiting list at time 0 the expected health gain discounted to that date is

$$q_t = \int_0^{w_t} h_t^o(s)e^{-rs}ds + (1-m_t)\int_{w_t}^{L_t^*(w_t)+w_t} h_t^*(s;w_t)e^{-rs}ds - \int_0^{L_t^o+w_t} h_t^o(s)e^{-rs}ds \qquad (80)$$

where we subscript time paths to allow for the possibility of technological change across periods. L_t^o, L_t^* are life expectancies without and with treatment measured from the date of treatment. Thus $L_t^o + w_t$, $L_t^* + w_t$ are life expectancies measured from the date placed on the list. r is the discount rate applied to health. Collecting terms

$$q_t = (1-m_t)\int_{w_t}^{L_t^*(w)+w_t} h_t^*(s;w_t)e^{-rs}ds - \int_{w_t}^{L_t^o+w_t} h_t^o(s)e^{-rs}ds = (1-m_t)q_t^*(w_t) - q_t^o(w_t) \quad (81)$$

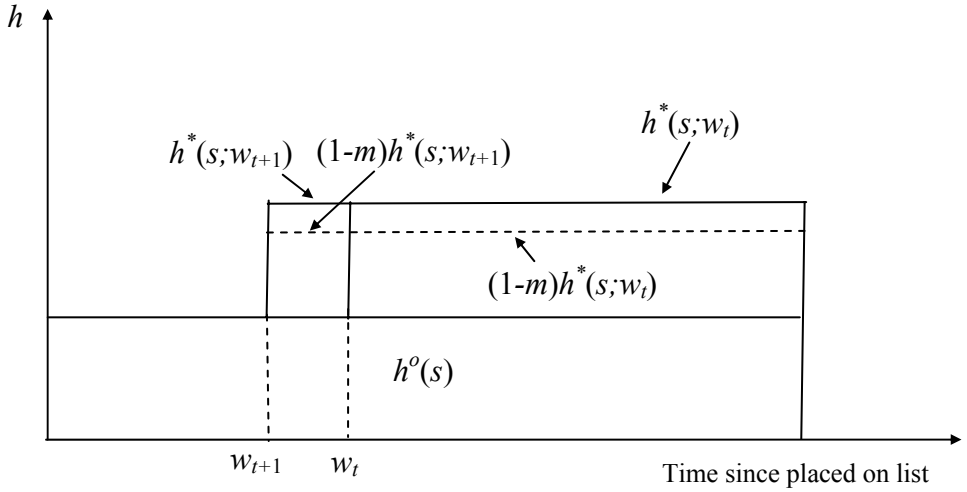
which is similar in form to the expression (26) for the health effect in the discussion of quality adjustment by mortality.

In general we will be unsuccessful in seeking a simple method of quality adjusting for changes in survival and in waiting times which is equivalent to the quality adjustment for survival derived in section 3.1, even with the assumption about unchanging technologies employed to construct the survival quality adjustment. The reason is that the survival probability enters the expression for health gain linearly but this is not true for the waiting time, as inspection of (80) reveals.

Only if further assumptions are made about the shape of the time paths of health status with and without treatment is it possible to derive a simple quality adjustment reflecting both survival and waiting times. The non-linearity of the waiting time adjustment also raises questions about whether the adjustment should be made on the basis of average waiting times applied to all cases of a particular type or whether we should use make the adjustment for each case separately and sum over the cases within each treatment type. We discuss this point further in section 4.10.4.

To get a reasonably simple expression for the waiting time adjustment, suppose the time paths h_t^o, h_t^* are constant with respect to elapsed time after the patient is placed on the waiting list, are unaffected by waiting time, do not change from one year to the next, and that total life expectancy is unaffected by treatment and waiting time: $L_t^o + w_t = L_t^*(w_t) + w$. (Remember that L_t^o, L_t^* are life expectancies without and with treatment measured from the date of treatment.) To simplify notation in this case where life expectancy from date of treatment is unaffected by treatment we use $L_t^o = L_t^* = L_t$. Figure 4.6 illustrates.

Figure 4.6 Health gain and waiting time – simple case



The health gain from treatment in Figure 4.6

$$q_{jt}^{1}(w_t) = \left(a_{jt}h_{jt}^{*} - h_{jt}^{o}\right)\int_{w_{jt}}^{w_{jt}+L_{jt}} e^{-rs}ds = \left(a_{jt}h_{jt}^{*} - h_{jt}^{o}\right)\frac{\left(e^{-rw_{jt}} - e^{-r(L_{jt}+w_{jt})}\right)}{r}$$

$$= \left(a_{jt}h_{jt}^{*} - h_{jt}^{o}\right)\frac{e^{-rw_{jt}}\left(1 - e^{-rL_{jt}}\right)}{r} \qquad (82)$$

Waiting has a cost, in units of health, which increases with the length of the wait but at a decreasing rate: an extra day after a long wait costs less than an extra delay after a short wait. The cost of a positive wait measured in terms of QALYs can be defined as this formula as

$$\kappa_t^{1}(w_t) = \left[q_t^{1}(0) - q_t^{1}(w_t)\right] = \left(\frac{a_t h_t^{*} - h_t^{o}}{r}\right)\left(1 - e^{-rw_t}\right)\left(1 - e^{-rL_t}\right) \qquad (83)$$

The discounted health effect (82) is convex in the waiting time: increases in the wait reduce the quality of care but do so at a decreasing rate. Thus the gain from a reduction in the waiting time from 40 weeks to 39 weeks is less than the gain from a reduction from 10 to 9 weeks. It has been suggested (Atkinson, 2005; 119) that the value of waiting time is more plausibly concave in the waiting time: the value of reductions is greater the longer the wait.

The convexity of the health effect in the waiting time implies that an increase in the dispersion of waiting times, holding the mean wait constant, increases the average value of treatment. This seems counter-intuitive.[6]

The quality adjustment factor is

$$\frac{q^1_{t+1}}{q^1_t} = \frac{\left[(1-m_{t+1}) - h^o/h^*\right]}{\left[(1-m_t) - h^o/h^*\right]} \frac{\left[e^{-rw_{t+1}}\left(1-e^{-rL_{t+1}}\right)\right]}{\left[e^{-rw_t}\left(1-e^{-rL_t}\right)\right]} \tag{84}$$

and the cost weighted index with a waiting time adjustment is

$$I^{xaw}_{ct} = \sum\nolimits_j (1+g_{xjt}) \left(\frac{a_{jt+1} - h^o_j/h^*_j}{a_{jt} - h^o_j/h^*_j}\right) \left(\frac{e^{-rw_{t+1}}\left(1-e^{-rL_{t+1}}\right)}{e^{-rw_t}\left(1-e^{-rL_t}\right)}\right) \omega^{jt}_{ct} \tag{85}$$

The use of the adjustment factor (84) requires estimates of life expectancy for patients. If we do not wish to use such estimates because we feel they are unreliable we can instead use

$$\frac{q_{t+1}}{q_t} = \frac{\left[(1-m_{t+1}) - h^o/h^*\right]}{\left[(1-m_t) - h^o/h^*\right]} e^{-r(w_{t+1}-w_t)} \tag{86}$$

$e^{-rw_{t+1}}/e^{-rw_t}$ will have the same sign as $\left[e^{-rw_{t+1}}\left(1-e^{-rL_{t+1}}\right)\right]/\left[e^{-rw_t}\left(1-e^{-rL_t}\right)\right]$. If life expectancy increases from t to t+1 then the adjustment using only $e^{-rw_{t+1}}/e^{-rw_t}$ will understate growth and will overstate growth if life expectancy falls. Using only $e^{-rw_{t+1}}/e^{-rw_t}$ and ignoring life expectancy is equivalent to assuming that life expectancy does not change.

Given the general lack of information on health with and without treatment we can use an estimate of the mean value of h^o_j/h^*_j (k) from a small sample of treatments (see

[6] Although the conclusion that a mean preserving increase in dispersion is welfare increasing is counter-intuitive, similar conclusions hold in other contexts. For example, a mean preserving spread in the distribution of prices of a good faced by a consumer who is risk neutral towards income will, if the consumer observes the price before buying the good, reduce expected utility. It is essential to specify the decision context carefully and in particular whether uncertainty is resolved before or after decisions are taken. In the case of consumer price uncertainty, it can be shown that if the consumer decides on consumption before the price is revealed and if she is averse to income risk then she is made worse off by a mean preserving increase in the dispersion of prices.

section 4.8.2). If we feel that this estimate is too unreliable we can set $h_j^o / h_j^* = 0$ in the index (85) and apply the adjustment for waiting with a pure survival adjustment.

Notice that we have applied the same discount rate to all types of treatment. The formulation of the health gain as having the same value whatever the type of treatment which underpins I_{ct}^{xaw} implies that the same discount rate should be applied to waits for heart bypass operations [HRG E04] as for varicose vein stripping [HRG Q11]. The differences in the health gains from the two treatments is already reflected in the index because it is in the cost weights which we have assumed are identical to the value weights. If we think that the assumption of proportionality of unit costs and marginal social values (30) is incorrect we could argue for using a higher discount rate for heart bypass waits than for varicose vein waits. However, this would be an ad hoc adjustment to a problem ($c_{jt} \neq \lambda_t p_{jt}$) which is better tackled more directly by adjusting the weights rather than the discount rates.

Because we treat waiting times as delaying health improvement we have also applied the same discount rate to waiting times and life expectancy. It is possible that people may feel differently about future health whilst waiting for treatment and when they have been treated and apply a discount rate whilst waiting for treatment of r_w but a rate of $r_L < r_w$ after treatment . The health effect would be

$$q_{jt}^{1'} = \int_0^{w_{jt}} h_{jt}^o e^{-r_w s} ds + \int_{w_{jt}}^{w_{jt}+L_{jt}} a_{jt} h_{jt}^* e^{-r_L s} ds - \int_0^{w_{jt}+L_{jt}} h^o e^{-r_L s} ds$$

$$= h_{jt}^o \left[\frac{(1 - e^{-r_w w_{jt}})}{r_w} - \frac{(1 - e^{-r_L w_{jt}})}{r_L} \right] + \left(a_{jt} h_{jt}^* - h_{jt}^o \right) \frac{e^{-r_L w_{jt}} \left(1 - e^{-r_L L_{jt}} \right)}{r_L} \quad (87)$$

4.10.2.2 Discounting to date of treatment with charge for waiting

We measure activity when it takes place which suggests that we should measure the benefit from treatment at the time it takes place. This is what we did in section 4.8.3 when discussing life expectancy and health effects. But with such assumption the discounted sum of QALYs, is, continuing with the simple assumptions about the time streams of health used in the previous section and in section 4.8.3, just

$$q_{jt} = a_{jt}\int_0^{L_{jt}} h_j^* e^{-rs} ds - \int_0^{L_{jt}} h_j^o e^{-rs} ds = \left(a_{jt}h_j^* - h_j^o\right)\left(\frac{1-e^{-rL_{jt}}}{r}\right) \qquad (53)$$

To capture the welfare lost as a result of having to wait for treatment we use a charge for waiting which is offset against (53) which captures only the benefit for the life span from date of treatment. As with any cumulating debt, interest is charged on the cost of waiting:

$$\kappa_{jt}^2 = \left(a_{jt}h_{jt}^* - h_{jt}^o\right)\int_0^{w_{jt}} e^{rs} ds = \frac{\left(a_{jt}h_{jt}^* - h_{jt}^o\right)}{r}\left(e^{rw_{jt}} - 1\right) \qquad (88)$$

The effect of treatment after adjustment for waiting time is now

$$\begin{aligned} q_{jt}^2 &= \left(a_{jt}h_{jt}^* - h_{jt}^o\right)\left[\int_0^{L_{jt}} e^{-rs} ds - \int_{-w_{jt}}^{0} e^{-rw_{jt}}\right] \\ &= \frac{(a_{jt}h_{jt}^* - h_{jt}^o)}{r}\left[2 - e^{-rL_{jt}} - e^{rw_{jt}}\right] \end{aligned} \qquad (89)$$

q_{jt}^2 is decreasing and convex in the waiting time so that an increase in an already long wait has a greater effect than the same increase in a short wait. Moreover an increase in the dispersion of waits reducing the expected value of treatment and there is a non-trivial cost of waiting even for the very young.

We can also allow for the possibility that the interest charge on waiting time differs from the interest rate on future QALYs:

$$\begin{aligned} q_{jt}^{2'} &= \left(a_{jt}h_{jt}^* - h_{jt}^o\right)\left[\int_0^{L_{jt}} e^{-r_L s} ds - \int_{-w_{jt}}^{0} e^{-r_w w_{jt}}\right] \\ &= \left(a_{jt}h_{jt}^* - h_{jt}^o\right)\left[\frac{\left(1-e^{-r_L L_{jt}}\right)}{r_L} - \frac{\left(e^{r_w w_{jt}} - 1\right)}{r_w}\right] \end{aligned} \qquad (90)$$

A difficulty with (89) is that it is theoretically possible for q_{jt}^2 to be very small or even negative (if $e^{rw_{jt}} > 2 - e^{-rL_{jt}}$) even if $a_{jt}h_{jt}^* - a_{jt}^o$ is positive. The same possibility arises with (90). However, this will only be a problem when the waiting time is similar to life expectancy. Table 4.5 has illustrative calculations for an example in which the health effects and life expectancy do not change from period *t* to period *t*+1illustrates for examples. Only in the three italicised cells where life expectancy and the waiting

times are similar do we observe nonsensical results where a reduction in waiting time leads to a very large or a negative adjustment. Otherwise the results are in line with intuition: the effect of a reduction in waiting time is greater at higher waits and for shorter life expectancy.

For all except the very shortest life expectancies a given percentage reduction in the waiting time implies a smaller percentage increase in quality adjusted output. The reason is that the waiting time is in general small relative to the length of time over which any health improvement will be enjoyed. The reductions are however in general not trivial. For example, reductions in waiting time from 120 to 90 days increases the growth rate by between 1.8% for life expectancy of 5 years and 0.35% for life expectancy of 30 years.

Table 4.5 (a) Waiting time adjustment q_{jt+1}/q_{jt}: discounting to date treated with a charge for waiting ($r_w = r_L = 1.5\%$, no change in health effects and life expectancy between period t and $t+1$)

Reduction in wait (days)	Life expectancy (years)						
	0.5	1	2	5	10	20	30
10 to 0	1.0582	1.0284	1.0141	1.0057	1.0030	1.0016	1.0011
30 to 10	1.1319	1.0602	1.0290	1.0116	1.0060	1.0032	1.0023
60 to 30	1.2469	1.0995	1.0456	1.0177	1.0090	1.0048	1.0034
90 to 60	1.3283	1.1106	1.0478	1.0180	1.0091	1.0048	1.0034
120 to 90	1.4897	1.1245	1.0503	1.0184	1.0092	1.0049	1.0035
150 to 120	1.9621	1.1424	1.0530	1.0188	1.0093	1.0049	1.0035
180 to 150	*27.2659*	1.1663	1.0561	1.0191	1.0094	1.0049	1.0035
365 to 150	*-0.1686*	*-38.6866*	1.6183	1.1563	1.0719	1.0366	1.0257

Table 4.5 (b) Waiting time adjustment q_{jt+1}/q_{jt}: discounting to date treated with a charge for waiting ($r_w = 10\%$, $r_L = 1.5\%$, no change in health effects and life expectancy between period t and $t+1$)

Reduction in wait (days)	Life expectancy (years)						
	0.5	1	2	5	10	20	30
10 to 0	1.0583	1.0284	1.0141	1.0057	1.0030	1.0016	1.0011
30 to 10	1.1326	1.0605	1.0292	1.0116	1.0060	1.0032	1.0023
60 to 30	1.2503	1.1006	1.0461	1.0179	1.0091	1.0049	1.0035
90 to 60	1.3376	1.1129	1.0488	1.0184	1.0093	1.0049	1.0035
120 to 90	1.5161	1.1285	1.0517	1.0189	1.0094	1.0050	1.0036
150 to 120	2.0850	1.1488	1.0550	1.0194	1.0096	1.0051	1.0036
180 to 150	*-10.6470*	1.1766	1.0587	1.0199	1.0098	1.0051	1.0036
365 to 150	*-0.1420*	*-9.6840*	1.6882	1.1679	1.0768	1.0390	1.0274

4.10.3 Optimal waiting times

All the waiting time adjustments considered above assume that a reduction in waiting time is always of value – shorter waiting times are better than longer. It has been suggested that to us that the assumption may not be valid for very short waiting times because some patients find treatment at very short notice to be inconvenient. It would be possible to conduct patient surveys or stated preference experiments to determine optimal waits and the costs associated with departures from them. But in the absence of such data we have to fall back on some crude assumptions which nevertheless enable us to determine how much impact such consideration might have. We have therefore investigated the implications of replacing the measure of waiting time w_{jt} in the waiting time adjustments considered in previous sections with $\hat{w}_{jt} = \min(w_{jt} - w^*, 0)$. This has the effect of increasing the proportionate effect of a given reduction in w_{jt} if w_{jt} exceeds w^* and reducing it to zero if $w_{jt} < w^*$. Another possibility would have been to define $\hat{w}_{jt} = (w_{jt} - w^*)^2$ but we felt that the implication that a reduction in w_{jt} below w^* reduced quality adjusted output was likely to prove difficult to justify.

4.10.4 Distribution of waiting times

We need to consider what waiting time measure should be used in the waiting time quality adjustments. Patients within an HRG do not have the same waiting times. This raises the question of how we take account of the variability of waiting times.

Variations in waiting times across patients for a given treatment are due to a combination of prioritization of patients so that different types have different expected waits and random factors so that patients of given types have uncertain waiting times. We will not be able to observe all the factors which are used in prioritisation so that the distributions of realised waiting times conditional on variables, such as age and gender, which are observable for individual patients in routine data, will overstate the amount of uncertainty for individual patients.

Suppose that treatment has the effect on health shown in the simple case in Figure 4.6. Then, assuming that all patients have the same life expectancy, the average health

gain for a patient, is discounting to the date of treatment as in section 4.10.2.2,

$$Eq_t^2(w) = \left[a_t h_t^* - h_t^o\right]\left[2 - e^{-rL_t} - Ee^{rw_t}\right]r^{-1} \tag{91}$$

where Ee^{rw} is the average of e^{rw} over realised *w*. Since $Ee^{rw} < e^{rEw}$ the waiting time adjustment factor $\left[2 - e^{-rL_t} - Ee^{rw_t}\right]r^{-1}$, using the mean wait to quality adjust the output for a particular HRG will lead to an overestimate of the mean quality adjusted output in a particular year. But what matters for the calculation of the output index is the ratio of the quality adjustment factors for consecutive years and in general

$$\frac{2 - e^{-rL_{t+1}} - e^{rEw_{t+1}}}{2 - e^{-rL_t} - e^{rEw_t}} \tag{92}$$

may be less than or greater than

$$\frac{2 - e^{-rL_{t+1}} - Ee^{rw_{t+1}}}{2 - e^{-rL_t} - Ee^{rw_t}} \tag{93}$$

There are two possible ways of allowing for the distribution of waiting times in the waiting time adjustment. We can write Ee^{rw} as

$$Ee^{rw} = \int_0^\infty e^{rw} f(w)dw = 1 + rEw + \frac{r^2E(w)^2}{2!} + \frac{r^3E(w)^3}{3!} + \cdots \tag{94}$$

or as the moment generating function for the distribution of waiting times. If the distribution is exponential (which it would be a random selection of patients was drawn from the waiting list each day) then $E_{w^i}e^{rw^i} = \mu_i/(1\text{-}r)$ where $1/\mu_i$ is the mean waiting time and the variance of waits is $1/\mu_i^2$. Similarly if the distribution is uniform on $[0,b_i]$ then $E_{w^i}e^{rw^i} = (e^{b_i r} - 1)/b_i r$. Thus if the distribution of waiting times has a tractable distribution and we know its parameters we can allow for the uncertainty in waiting times without requiring individual level data. After inspection of the empirical distributions so we decided not to pursue this approach as there is considerable disparity in the empirical distributions and few seem to be approximated by tractable theoretical distributions.

If individuals have uncertain waiting times and have a utility function which is decreasing and convex in waiting time then taking the mean of the actual waiting times will understate the costs of waiting. One way to allow for a cost of the risk of having very long waits is to not to use the mean or median of the distribution of waits

but the "certainty equivalent" wait w^c : the mean wait plus a waiting time "risk premium". The certainty equivalent waiting time is

$$EU(w) = u(w^c) = u(E(w) + RP) \quad (95)$$

where *RP* is risk premium and *u*(.) is a utility function.[7] We take the specification of the benefits from treatment when there is a positive waiting time in section 4.10.2.2 as the utility function. The standard approximation to the risk premium (Pratt 1964) is

$$RP = -A\sigma_w^2 / 2 \quad (96)$$

where $A = -u''/u'$ is the coefficient of absolute risk aversion and σ_w^2 the variance of waiting times. With the specification in section 4.10.2.2, the coefficient of absolute risk aversion is $-r$ so that

$$RP = r\sigma_w^2 / 2 \quad (97)$$

Table 4.6 shows, for different rates of discount, the proportion of elective cases admitted in 2002/3 in HRGs where the distribution of waiting times implies a certainty equivalent wait greater than various percentiles of the distribution.

[7] Since *u* is decreasing in the waiting time we have defined the risk premium as an addition to the mean by contrast to the usual case where *u* is increasing in a risky variable such as income. Hence the minus sign in the definition used here.

Table 4.6 Number of elective HRGs and proportion of total elective admissions in HRGs where the certainty equivalent waiting time w^c exceeds particular percentiles of the distribution of waiting times in 2002/3

Certainty equivalent wait greater than	70th	75th	80th	85th	90th
r = 0.01					
Number HRGs	214	152	152	95	32
Prop cases	0.109	0.057	0.057	0.032	0.005
r = 0.015					
Number HRGs	222	156	156	97	32
Prop cases	0.138	0.069	0.069	0.038	0.005
r = 0.03					
Number HRGs	241	170	170	107	34
Prop cases	0.235	0.153	0.153	0.116	0.005
r= 0.05					
Number HRGs	258	185	185	112	38
Prop cases	0.292	0.205	0.205	0.123	0.006
r = 0.1					
Number HRGs	292	211	211	127	42
Prop cases	0.314	0.224	0.224	0.130	0.008
r = 0.15					
Number HRGs	320	233	233	144	52
Prop cases	0.375	0.241	0.241	0.141	0.013
r = 0.2					
Number HRGs	340	250	250	154	56
Prop cases	0.392	0.258	0.258	0.142	0.014

The table suggests that using the extreme top end of the distributions as certainty equivalent waiting times is likely to overstate the cost of risk arising from the distribution of waiting times. In our calculation of the indices with waiting time adjustments in section 5 and 6 we compare the effects of using the 80th percentiles as the certainty equivalent wait compared with the mean wait at rates of interest between 1.5% and 10%

The alternative approach is to use individual level data from HES on the waits experienced by each elective patient. Given the number of electives this is fairly cumbersome since each calculation of an index takes around four hours with individual data. We do however report in section 5 the results of some calculations of indices to show the sensitivity of indices to the use of individual rather than a certainty equivalent wait.

4.10.5 Outpatient waits

Table 4.7 highlights the mean waiting times in weeks for a select number of key specialties. In most specialties waiting times have fallen.

Table 4.7 Outpatient waiting times in weeks, selected specialties, 1999/00-2003/04

Code	Specialty	1999/00	2000/01	2001/02	2002/03	2003/04
100	General Surgery	7.448	7.398	7.495	6.787	6.377
120	ENT	11.798	11.261	10.764	9.871	9.258
140	Oral Surgery	9.789	9.451	9.707	9.029	8.653
150	Neurosurgery	10.546	10.781	10.812	10.019	9.251
160	Plastic Surgery	12.370	12.352	13.023	10.300	9.233
170	Cardiothoracic Surgery	3.865	4.162	4.150	4.194	4.503
300	General Medicine	8.799	8.877	9.066	8.099	7.392
310	Audiological Medicine	13.815	13.175	13.813	10.986	9.351
340	Respiratory Medicine	6.580	6.986	7.415	6.953	6.369
350	Infectious Diseases	5.978	6.491	6.583	5.232	4.949
360	Genito-Urinary Medicine	2.342	2.316	2.483	2.214	2.048
370	Medical Oncology	3.535	3.597	3.208	3.559	3.851
400	Neurology	13.072	12.814	13.378	11.324	10.038
420	Paediatrics	6.680	6.758	6.756	6.639	6.582
430	Geriatric Medicine	4.975	5.170	5.344	5.340	5.232

Ideally, because of the non-linearity of the effect of waiting time on the value of activity, we require information on the total time that each individual waits for treatment. However, individual HES records of inpatient waiting times are not linked to individual outpatient data and indeed outpatient data is available only aggregated to specialty level which cannot be linked satisfactorily to inpatient data aggregated to HRG level.

If we wanted to discount health effects to date of treatment (as in section 4.10.2) we require an estimate of $e^{rw_{\ell t}}$ where $w_{\ell t}$ is individual ℓ's total wait for treatment which is the sum of her outpatient and inpatient wait: $w^O_{\ell t} + w^I_{\ell t}$. We do not have such linked individual level outpatient and inpatient data. If we could group outpatient data to the same HRGs or specialties as inpatient data we could use the average wait for outpatient treatment for each group to calculate $e^{r(w^O_t + w^I_t)}$ for a specialty or HRG. But this is not feasible. In the future it will be possible to use HES to get total waiting

times for individuals across outpatient and inpatient care. But for the moment we require a means of incorporating a quality adjustment for outpatient waits.

There is data on first outpatient visits, follow up visits, HRG outpatients, and maternity outpatients but only data on the wait for the first visit. The data on the HRG outpatient and maternity outpatients do not have speciality codes that we can match to the speciality codes for the first and follow up visits. We propose to use the waiting time adjustments from section 4.10.2.2 but to apply the waiting time for first visits to both first and follow up visits. It can be argued that a reduced first outpatient visit wait reduces the delay until the value of the whole course of visits is realised. Second, less plausibly, it could be argued that, regarding first and follow up visits as having different values, reductions in waits for first visits also indicate that waits for follow up visits have fallen and that this increases the value of follow up visits as well.

4.10.6 Waiting time adjustment: conclusion

In this sub section we have

- shown how data on waiting times can be used to quality adjust the cost weighted output index in addition to the survival based adjustments considered in sections 4.7 and 4.8.
    - regarding waiting time as a characteristic separate from the health effect of treatment, yielding an adjustment which is additive the health effects
    - regarding waiting time as delaying the health effect of treatment, yielding adjustments which act as scaling factors on the health effects
- concluded that calculation of an index based on the value of the health and waiting time characteristics, as in section 2.4 and section 4.10.1, is not currently possible because of the lack of health effects data. We have constructed waiting time quality adjustments which can be applied to all elective HRGs with existing data and which rest on different assumptions about whether one should discount to the date the patient is placed on the waiting list (section 4.10.2.1) or to the date of treatment (4.10.2.2,).
- suggested that an adjustment based on discounting to the date of treatment are preferable since they imply that increased dispersion of waiting times would

reduce quality adjusted output, whereas discounting to date placed on the list implies that patients would prefer increased dispersion.

- suggested that the dispersion of waiting times should be reflected in the waiting times adjustment by calculating the adjustment using, not the mean or median wait, but the 80th percentile waiting time which can be interpreted as an estimate of the "certainty equivalent" wait.

As with the health effects adjustments a final view on the preferred adjustment depends on the plausibility and lack of volatility of the adjustments when confronted with the data in sections 5 and 6.

Table 4.8 Summary of waiting time quality adjustments

Quality adjustment	Wght	Form	Rationale	Results	Assumptions	Comments
Survival. Uniform health effect. Life expectancy. Discounting to date on list.	Cost	$\frac{\sum_j c_{jt} x_{jt+1}\left(\frac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)\left(\frac{e^{-rw_{jt+1}}\left(1-e^{-rL_{jt+1}}\right)}{e^{-rw_{jt}}\left(1-e^{-rL_{jt}}\right)}\right)}{\sum_j c_{jt} x_{jt}}$	4.10.2.1	Sec 5.5, Sec 6.3. With "optimal" wait adjustment Sec 5.5	Efficient allocation. Constant health status. No effect of treatment on life expectancy. k_j equal for all j in Sec 5.5. k_j varying across j in secs 6, 7	Adjustments have more effect with more costly treatments. Also calculated with different discount rates for w and L
Survival. Uniform health effect. Life expectancy. Discounting to date treated, with charge for wait.	Cost	$\frac{\sum_j c_{jt} x_{jt+1}\left(\frac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)\left(\frac{2-e^{rw_{jt+1}}-e^{-rL_{jt+1}}}{2-e^{rw_{jt}}-e^{-rL_{jt}}}\right)}{\sum_j c_{jt} x_{jt}}$	4.10.2.2	Sec 5.5, Sec 6.3	Efficient allocation. Constant health status. No effect of treatment on life expectancy. k_j equal for all j in Sec 5.5. k_j varying across j in secs 6, 7	Adjustments have more effect with more costly treatments. Also calculated with different discount rates for w and L
Health and waiting time characteristics	Value	$\frac{\sum_j x_{jt+1}\left[\pi_{ht}\left(a_{jt+1}h_j^* - h_j^o\right)\left(1-e^{-rL_{jt+1}}\right)r^{-1} - \pi_{wt} w_{jt+1}\right]}{\sum_j x_{jt}\left[\pi_{ht}\left(a_{jt}h_j^* - h_j^o\right)\left(1-e^{-rL_{jt}}\right)r^{-1} - \pi_{wt} w_{jt}\right]}$	2.7, 4.10.1	Sec 6.5	Constant health status. No effect of treatment on life expectancy. Value additive in health and waiting time.	

c_{jt} unit cost, volume x_{jt} volume; a_{jt}. m_{jt} proportion patients alive, dead on discharge (or after 30 days); ; k estimate of proportionate effect of treatment on quality adjusted life years (QALYs); L_{jt} life expectancy at mean age of patients treated; r discount rate on QALYs/waits; π_{ht} value of QALY (£s), π_{wt} value of day's wait, h^*, h^o estimates of health with and without treatment.

4.11 Patient satisfaction

The indices we have set out above are all variants on cost weighted aggregates built up from FCE data. There are however some attributes which the Department may wish to see reflected in an output index which are not covered by the data discussed so far. Prominent among these are patient views on cleanliness, food quality and satisfaction with non-curative aspects of treatment (such as the behaviour of nurses and doctors).

Our analysis in this section is based on the assumption that these variables cause patient disutility because they are seen as indicative of poor treatment quality. For this part of our work then, we introduce these to our index is by identifying indicators (such as aggregates constructed from responses to surveys of patient opinion). We then calculate the growth rate of our comprehensive index, I_t^{Comp} as the weighted sum of the growth rate of one of the quality adjusted output indices we have described earlier (call it I_t^0) and the growth rates of the other indicators. We denote the growth rate of indicator k by I_t^k. If there are n such indicators, and the relevant weights are denoted by ω_k then the overall index is given by

$$I_t^{Comp} = \sum_{k=0}^{n} \omega_k I_t^k \qquad (98)$$

In some circumstances we may be able to deduce appropriate values for the ω_k. For example we know the total value of expenditure on cleaning and this therefore gives us the basis for a weight on a cleanliness measure. In other cases the choice has to be purely arbitrary. We do not know what costs, if any, are involved in persuading doctors and nurses to be polite to patients, so there is little we can do except to invent a weight.

The NHS Patient Survey Programme covers Acute Trusts, Primary Care Trusts, Mental Health, Ambulance Trusts and others. In addition, there are plans for surveys focusing on the National Service Frameworks for coronary heart disease, mental health, older people, diabetes, etc. There are a number of issues that make the

incorporation of patient satisfaction into our measure of NHS output difficult, some theoretical and some practical.

From a theoretical standpoint, there are at least three reasons why measures of patient satisfaction should be treated with caution. The first is that patient satisfaction is not independent of the level of output, x_j, or other characteristics q_k. For example, if we include a measure of waiting times in our set of characteristics, it is likely that patient satisfaction, not only with the time they had to wait, but also other aspects of the service they have received, will be correlated with this. Because of this, there is the likelihood of 'double-counting' output.

Second, patients' reports of satisfaction will depend on the levels of service they expect. Certain sections of the population may have lower expectations than others, as noted in the First Interim Report (Dawson *et al*., 2004a). These expectations may vary systematically across the population, introducing bias to our measure of output.

Third, satisfaction is a multidimensional concept. There may be a number of orthogonal aspects to satisfaction, beyond total satisfaction, that should be considered as separate characteristics in themselves and aggregated using their social valuations. For example, patients' satisfaction with the hotel services aspect of an inpatient stay may be largely independent of their satisfaction with the medical aspects. Eliciting these different dimensions of satisfaction with single instruments is extremely difficult, if not practically impossible, which leads us to the empirical obstacles to the inclusion of a measure of patient satisfaction into the index.

The practical question arising from the use of patient satisfaction data is the construction of appropriate growth indicators, I_t^k In our analysis we make use of qualitative surveys carried out for the Health-care Commission. The results of these surveys show what proportion of patients gave answers in particular categories to each question.

We constructed aggregates using the same approach as the Health-care Commission. For each relevant question (i.e. those questions which were asked in more than one survey round without changes likely to influence results), we gave a weight of 100 to the most favourable answer, a rate of 0 to the least favourable answer and weights to

intermediate categories based on the number of possible choices. Thus if there was one intermediate category it was given a weight of 50, if there were two the more favourable was given a weight of 67 and the less favourable a weight of 33 and so on. These weights were applied to the proportion of respondents in each category in order to give us an overall score. For a group of overall questions (cleanliness, food quality or other aspects of satisfaction) we took the arithmetic average of the individual scores in order to produce an overall indicator.

This quantification appears arbitrary and one might be concerned that it does not really identify the latent variables which underlie the reported categorical responses. However the indices produced in this simple manner were very strongly correlated with indices derived from assumptions about the density function of the underlying latent variables. Thus we are reasonably happy with their use.

There are a number of questions we have not used in our analysis. Some of these collected information in both years but did not invite observations on service quality (e.g. What age group do you belong to?). Others we have rejected because there were small but possibly important changes to the questions between the years (e.g. in 2002 "Were you in a mixed-sex room or ward?" and in 2004 "Were you in a mixed-sex bay in a room or ward?", or because they do not actually convey what is needed to assess patient quality. Thus the question about the time people have waited to see GPs does not identify patients who have been unable to see their GPs because the GPs only make appointments for the day that they will see the patients.

The outpatient questionnaire includes one question about the time the patient has had to wait for an appointment once they knew they needed it. We have counted waiting time elsewhere and therefore leave that question out of our assessment. A second question asks how long the patient has waited relative to the time of their appointment, inviting categorical answers. By making the assumption that the mid-point of each category is relevant to all patients in that category and a plausible assumption for the final category (two and a half hours for a wait of more than two hours) we can estimate a mean waiting time and measure the percentage change in this. This can be combined with other score variables for outpatient treatment.

These indicators are measured on a per-patient basis. To give an overall change in the "volume of quality", it is necessary to add to each indicator the growth in the volume of treatment. Our own view is that this is not measured exactly by the cost weighted index, since non-medical quality matters whether a patient is given a cheap or an expensive treatment. The measure we use is number of inpatient spells. For outpatients the appropriate volume indicator is the number of outpatient consultations and the same is true for primary care. We discuss this further in section 5. The percentage change in the volume component has to be added to the percentage change in quality to give the change in the overall volume of quality, i.e. the relevant I_t^k.

The quantified data, for the relevant questions are shown in tables A1 to A4 in Appendix A.

4.12 Discount rate on health

There is little agreement on the appropriate discount rate to use when health gains accrue over time. Cairns (1994, 1997) examines evidence on individuals' time preferences for social health gains, such as saving lives of other people at different points in the future. Discount rates for saving lives fall from 41% where the delay is only two years to 16% where the delay is nineteen years. Gravelle and Smith (2001), using a social welfare framework, point out that if the value of health is increasing over time, estimates of the volume of health benefits must take this into account. The volume of health effects can be adjusted directly by the rate of growth in the value of health. An indirect method is to reduce the discount rate on health effects relative to the discount rate applied to costs. The size of the adjustment depends on estimates of the rate of growth of the value of health which itself is a function of the rate of growth of the direct utility effect of health and the rate of growth of income. If the indirect method of allowing for growth in the value of health by means of a lower discount rate on health effects is used, the discount rate on health effects will be 1-3% less than the discount rate on costs depending on the assumptions made about key variables. The Department of Health currently discounts health benefits at 1.5% p.a. and costs at 3.5% p.a. (Department of Health, 1996;, HM Treasury, 2003). We take 1.5% as our base case in discounting life expectancy but also consider higher rates to investigate

the sensitivity of the index to the discount rate.

The formulations for the waiting time adjustment allow for the possibility that different discount rates could be applied to the period whilst waiting and the post treatment period. It could be argued that a higher rate of discount should be applied to the waiting period since there is a cost to waiting over and above the delay in getting treatment. Unlike the discount rate applied to health effects post treatment there is no obvious higher rate to use for waits. We suggest that the direct disutility from waiting, over and above the delay in getting treatment, is best dealt with by treating it as a separate characteristic. Thus when the data become available to support a value weighted index waiting time could be included both as a determinant of the health effect and as a separate characteristic if patient preference studies suggest that this is appropriate.

We have included in Sections 5 and 6 illustrative calculations of the scaling effect of waiting time using different discount rates for health and waiting time. Since the rates applied to waiting time are arbitrary and in the absence of any contrary indication we prefer on balance to use the same rate for waits and health effects.

4.13 Quality adjustment for general practice

Until recently routine data on the quality of general practice has been virtually non-existent because information on general practice has been collected with the aim of paying GPs and until April 2004 the General Medical Services contract paid little explicit attention to quality. GPs received bonus payments linked to the proportion of eligible women receiving cervical screens, and of children vaccinated. Cervical screening activities were included in the old Cost Weighted Activity Index as volume measures of activity so that to this extent the CWAI had a small quality component. From 1998/9 onwards GPs have been able to opt for Primary Medical Service contracts negotiated with their Primary Care Trusts under which they are meant to deliver an enhanced package of services. By April 2004 around 35% of practices had switched to PMS. Since PMS contracts are locally negotiated, the central reporting of the activities targeted under the GMS contract is patchy.

There are a number of validated measures of general practice quality - such as the proportion of patients with coronary heart disease whose blood pressure is controlled. The Department of Health is investigating the use of the QRESEARCH database extracted from a large (around 500) sample of GP electronic record systems to measure the quality of care in general practice using these types of indicators (Simkins, 2005).

The new General Medical Services contract of April 2004 contained financial incentives linked to achievement against a large basket of quality indicators in the Quality and Outcome Framework (QOF) (Department of Health, 2003a; Roland, 2004). PMS practices were also required to join the QOF and returns for all practices are centrally collected and available via the Quality, Prevalence and Indicator Database QPID run by the DH's Health and Social Care Information Centre.

It is possible to calculate a subset of the quality indicators in the QOF for the practices in the QRESEARCH database over a number of years prior to the introduction of the QOF for all practices in April 2004. This series can then used to quality adjust the past volume series of general practice consultations.

In principle the quality indicator series can be calculated in future years for the practices in QRESEARCH and, for all practices. There are potential difficulties in constructing a quality adjusted series which covers the periods before and after 2004/5. GPs are partially altruistic: their actions are guided by a professional concern for their patient's well being and by a concern for their income and effort. The introduction of the QOF in April 2004 changed the relative importance of the professional and personal incentives. For aspects of quality covered by the QOF GPs now have both financial and professional incentives. For aspects of quality which are outside the QOF they now only have professional incentives. The QOF may therefore lead to increased activity in the areas covered by the QOF and reduced activity in those outside it. Hence the changes in QOF activity may provide a misleading guide to changes in overall quality of care since the unremunerated and therefore not centrally notified activities may have declined, or not increased to the same extent.

The data produced by the QOF is likely to be a fruitful source of quality adjustment of general practice activity in future years. Some of the indicators can be translated into health impacts, for example from studies of CHD event risk factors including blood pressure. But the interpretation of trends in quality indicators will require empirical modelling of the responses of practices to changing financial incentives and their impact on unremunerated quality indicators. Such modelling should be possible using QRESEARCH or a similar database.

The QOF also has an incentive for GPs to undertake surveys of their patients, though the reward is for carrying it out and acting on the results, and is not linked to the actual patient responses. If the QOF is adjusted to link payment to responses, for example on satisfaction with particular aspects of the practice then this will be a potentially fruitful source of data on the patient experience in general practice. We discuss in section 4.11 how such data could be used as a quality adjustment to an output index.

It has been suggested that admission rates for ambulatory care sensitive conditions (ACSCs) can be used as a measure of the quality of general practice. ACSCs are conditions which can be controlled in a good quality general practice and which should not result in a hospital admission. The usual examples include asthma, diabetes and epilepsy and ACSCs for these conditions have been used by the DH and the Commission for Health Improvement as primary care performance indicators. They have also been used extensively in the US, New Zealand and Australia (Giuffrida *et al.*, 1999; Jackson and Tobias, 2001; Victorian Government, 2004). Even if ACSC admission rates affected by the quality of general they are also strongly influenced by factors outside the control of GPs, including the availability and admission criteria of hospitals (Giuffrida *et al.*, 1999). ACSCs admissions are an imperfect proxy for the health of the relevant population. We do not recommend their use as a means of quality adjusting measures of national general practice output.

4.14 Atkinson principles and quality adjustment

The Atkinson Review published its final report in January 2005 setting out recommendations for improving the measurement of government output and productivity (Atkinson, 2005).

Atkinson paid particular attention to the Eurostat (2001) Handbook on price and volume measures in National Accounts. The Handbook made important recommendations on the methods to be used to measure output and the implications for the measurement of non-market output.

Eurostat distinguished between activities, outputs, and outcomes. For purposes of national accounting it is preferable to measure outputs (treatment received by a patient) rather than activities (number of operations or prescriptions). Outcomes should be used to quality adjust outputs. Eurostat "graded" the methods that governments use for measuring non-market outputs into A, B or C.

A. Preferred method:
    i) use output indicators (rather than activities)
    ii) all services should be covered, as detailed as possible
    iii) outputs should be quality adjusted
    iv) outputs should be cost weighted

B. Less satisfactory but acceptable method:
    i) use output indicators but the detail needs to be improved
    ii) no account is taken of quality change

C. Unacceptable method:
    i) use of inputs instead of activities or outputs
    ii) coverage of output not representative

In discussing Group A methods, counting inpatient activity by DRG is accepted as a way of measuring output but this appears inconsistent as it is only a more disaggregated way of counting activities. Atkinson acknowledges the difficulty of

quality adjusting by DRG activity. Atkinson says that ideally output should be measured as the whole course of treatment for an illness rather than just counting activities (by DRG or HRG). This would make it possible to take better account of the quality of care. The report notes that it "would be very helpful to be able to base quality adjustments for NHS output on a data set which measures the health outcome achieved as a result of treatment, collected annually by all or part of the NHS for most aspects of health care." In the short-term attention could be focused on a few disease groups for which data on outcomes is available.

In the general discussion of methodology, Atkinson points to the distortion in output indices that are weighted by average costs. Different outputs should be weighted by the marginal value of the outputs to individuals and there is no reason to believe this corresponds to marginal or average cost.

In major respects Atkinson (2005) recommended a methodology for measuring NHS output growth advocated in our earlier reports:

- Units of output should relate to patient journeys (courses of treatment) rather than activities (tests, procedures, consultations, drugs prescribed). Until a patient identifier permits linking the various services delivered to a single patient, output will still have to be measured as the sum of activities.
- The quality of outputs should be incorporated into output indices.
- Quality measures should reflect the attributes of output valued by individuals. Atkinson initially identifies relevant attributes as those recognised by current objectives of government health policy. We suggest, since individuals may value other attributes than those targeted by the policies of any particular government, and may value them differently, that further research is required on the attributes valued by individuals.
- In order to generate a single index of output, weights must be attached to the multitude of NHS activities. Weights should reflect the marginal social value of the activities. At present relevant data does not exist and ONS will continue to use cost weights. Atkinson acknowledges the distortions introduced by the use of cost weights—a relative increase in expensive treatments appears to increase NHS output while a relative increase in cost reducing treatments with

the same or better health outcomes will appear to reduce NHS output.

If we had routinely collected data on health outcomes and these data were used to weight the activities of the NHS, it would not be necessary to make a separate quality adjustment to an index of NHS output. In the short term Atkinson suggests an alternative approach. Changes in quality which are currently measurable should be used to augment cost weighted output indices. We suggest that for health care, currently available data permit quality adjustment in respect of two attributes: waiting times and survival rates. Following the Atkinson (2005) approach requires that we find a basis for determining the relative value of changes in these two attributes of health care and use the resulting estimate of the rate of change in quality to quality adjust the cost weighted output index.

Atkinson (2005) stresses the need to measure the value added by public services. This is particularly important in health care where individuals may be healthier or living longer for reasons unrelated to the availability and quality of health care. What is required are measures of the marginal effect of the NHS in producing outputs of value to patients. Given the difficulties in estimating production functions, rough and ready judgements will be required when attributing improvement in health outcomes to health care.

Atkinson (2005) discusses the "complementarity" between public and private output. As economic growth leads to higher standards of living, the relative valuation of different goods and services will change. This will affect the weights to be attached to the various activities in an output index. If people are living longer the value of a hip replacement may increase as the present value of the benefits are estimated over a longer period of time. We argue that such changes in value arising from factors outside the control of the NHS should not be counted in measuring the quality adjusted growth rates of different types of NHS output. Instead they should be included in the weights to be applied in calculating the weighted average rate of growth of NHS output (the output index). Changes in the value of a hip replacement arising from factors outside the control of the NHS should influence decisions on how to allocate NHS resources across different activities but not in assessing the rate of growth of the output of hip replacement activities.

5 Experimental indices of NHS output

5.1 General trends, index form and data sources

This section demonstrates the impact on a cost weighted output index of adjustments for a number of quality dimensions. Prior to presenting results, however, it is useful to highlight some salient features of the data.

Our starting point is the dataset used by the DH in its cost weighted output index (CWOI). For inpatient hospital care, the Hospital Episode Statistics (HES) are used (data on non-elective, elective and day case activity). The main quality adjustments discussed in this report are for survival and waiting times. Mortality rates are only available for HES activity (non-elective, elective and day case hospital activity). Waiting times are available for electives and day cases and some outpatient activities.

In 2002/03 the data consisted of 1913 groups of activities. Table 5.1 gives a summary of the main divisions of this dataset. In this chapter we focus on the quality adjustments to the HES data, which apply to 47% of the activity currently included in the CWOI. In chapter 6 we examine the effects of different quality adjustments to output indices for our small specimen set of HRGs for which we have health effects data. In the light of the results in this chapter for the HES set of hospital activities and the specimen set in chapter 6, we set out our preferred quality adjustment variant in chapter 7, for the set of HES hospital activities, for a broader set of hospital activities including outpatient treatments, accident and emergency, and for all NHS activity. In chapter 9, the output indices from chapter 7 are combined with the input indices from chapter 8 to yields estimates of various types of productivity growth.

Table 5.1 Activities and cost shares 2002/3

	Number of activities (millions)	Cost shares[1]
Electives+ day cases	5.58	13.38
Non-electives	5.96	22.1
Outpatients	53.43	11.15
Other activities[2]		53.37
Total		100

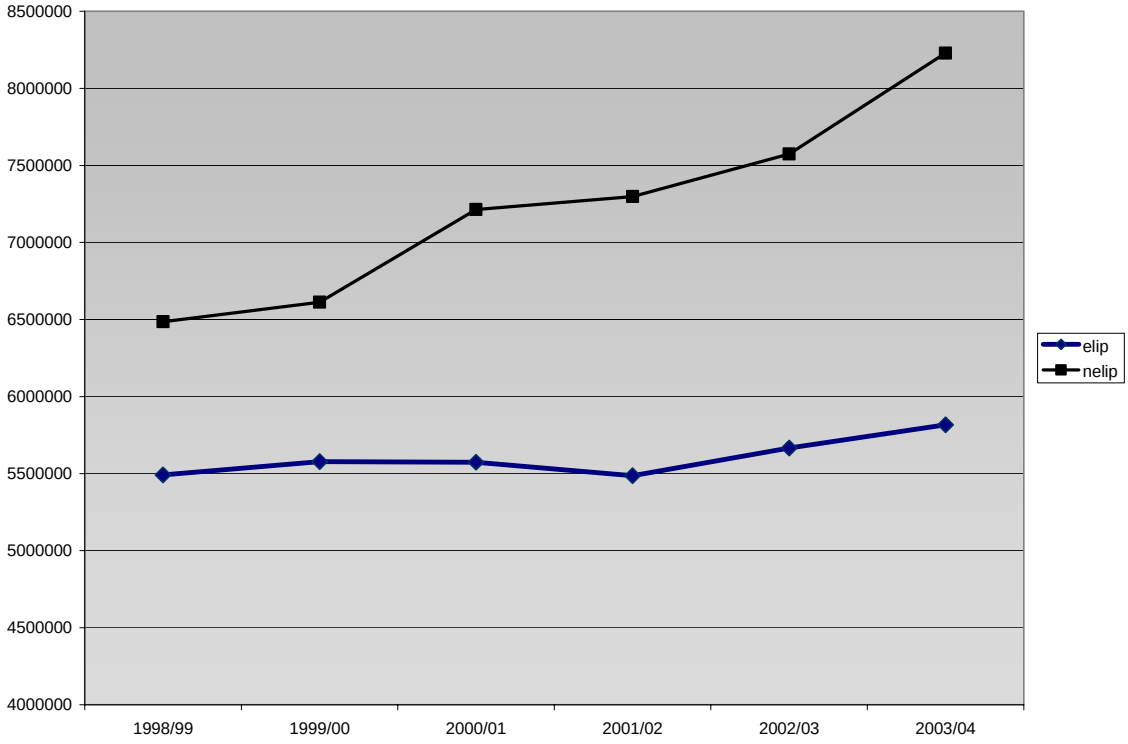
Notes: 1. Derived by multiplying activities by unit costs; 2.These activities are measured in non-comparable units so total numbers of activities are meaningless. A division of cost shares with this category is shown in Table 3.1.

In order to highlight the impact of quality adjustment for the activities where the data permit adjustment, in sections 5.4 and 5.5 we compare our quality adjusted indices to an unadjusted index restricted to the same set of activities. In the tables this truncated version of the CWOI is labelled "unadjusted". In Section 7 we examine how quality adjusting for this subset of activities affects the value of the complete CWOI.

5.1 General trends and data

The HES data are grouped according to procedures comprising 574 Healthcare Resource Groups (HRGs), with an additional separation into electives and day cases and non-electives (Appendix B). Figure 5.1 graphs the number of episodes for each year from 1998/99 to 2003/04. It shows little change in electives up to 2001/02 with some growth thereafter. Non-electives show more significant growth, with very high growth in the final year.

Figure 5.1 Number of FCEs, electives+day cases (elip) and non-electives (nelip), 1998/99 – 2003/04



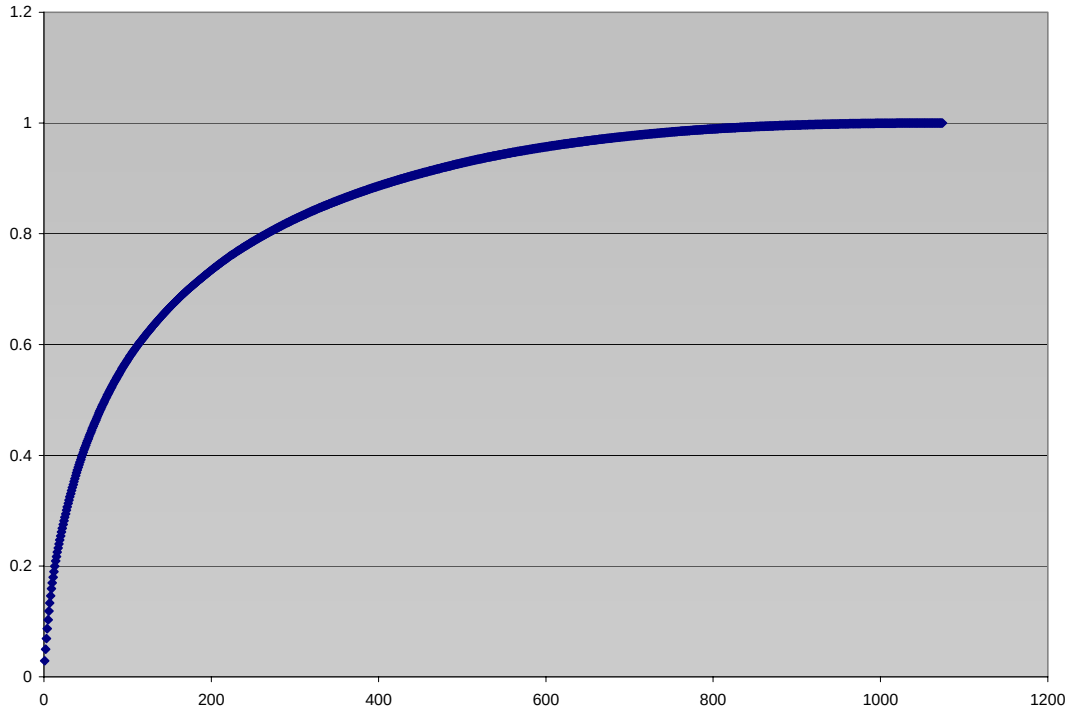
Within these broad categories there is considerable variation in number of procedures and in growth by HRG. For example, comparing 2002/03 with 2001/02, the arithmetic mean growth in episodes for elective HRGs was 3.8% but with a standard deviation of 15.9% – growth across HRGs was even more variable for non-electives. Partly this reflects substitution across treatments but nevertheless the variation is large.

Unit costs from the Reference Costs database are employed to aggregate these diverse activities. The unit costs also show considerable variation across procedures, from over £20,000 for transplant procedures to under £500 for ophthalmic and ear procedures. In 2002/03 the mean unit cost across HRGs for electives was about £1,700 with a standard deviation of £2,220, with figures of £2,200 and £2,600, respectively, for non-electives.

The cost weighted output index combines activity growth by weighting by unit costs, equivalent to multiplying the ratio of activities by cost shares. Cost shares are concentrated in a few HRGs. Treating electives and non-electives as separate sets of

activities, in 2002/03, 25% of expenditure was accounted for by only 20 HRGs with 50% accounted for by 74 HRGs, as illustrated by the cumulative expenditure share chart below. The top expenditure categories include HRGs where activity rates are very high such as maternity care for normal deliveries or hip replacements and which are not typically life threatening. But it also includes HRGs where mortality rates are very high such as heart procedures and complex procedures involving the elderly. A high cost share on this latter group turns out to be important in the adjustments for survival discussed below.

Figure 5.2 Cumulative expenditure shares, FCEs (537 elective, 537 non-elective HRGs)



5.2 Index form

Table 5.2 compares the Laspeyres (base period weighted) index with the Paasche (current period weights) index and the Fisher index which is the geometric mean of the Laspeyres and Paasche indices. The index number formula used does have a small

but significant impact on the indices, as shown for episodes for selected years. In this report we follow ONS in reporting Laspeyres indices.

Table 5.2 Impact of index number formula on CWOI index, FCEs based

	1999/00-2000/01	2000/01-2001/02	2001/02-2002/03
Laspeyres	0.90%	0.93%	4.41%
Paasche	0.71%	0.82%	4.47%
Fisher	0.81%	0.87%	4.44%

Note: the reference costs for 1998/99 were considered unreliable so the index for the comparison between 1998/99 and 1999/00 use 1999/00 unit costs in the numbers reported below. 2003/04 unit reference costs on a comparable basis were not available.

5.3 Spells versus episodes

We discussed the choice between measuring hospital output in finished consultant episodes (FCEs) and continuous inpatient spells (CIPS) which consist of sets of consecutive FCEs in section 4.2 (see also Appendix B). We argued that CIPS were a better approximation to the patient journey and therefore a more appropriate measure of output. We use CIPS for our calculation of the effects of quality adjustments.

Although there are around 8% fewer CIPS than FCEs this should have essentially no effect on the calculation of a cost weighted output index since we constructed our unit costs for CIPS from the underlying FCE unit costs. Table 5.3 compares FCE based and CIPS based CWOIs as a check on our calculations of unit costs of spells. The only reason for a divergence between the two indices is that some of the FCEs assigned to a particular year in the FCE index may be assigned to a different year in a CIPS index since a CIPS is assigned to a year only if its last FCE finished in that year.

We would however expect to see differences in FCE and CIPS based indices once the outputs are adjusted for survival and mortality since these adjustments are applied to the different distributions of HRG types generated by the FCE and CIPS volume measures.

Table 5.3 Comparison of cost weighted output indices for hospitals based on finished consultant episodes and continuous inpatient spells

	CWOI index		
	Episodes	CIPS	pp diff
1998/99-1999/00	1.84	1.87	-0.03
1999/00-2000/01	0.90	0.91	-0.01
2000/01-2001/02	0.93	0.95	-0.02
2001/02-2002/03	4.41	4.44	-0.03
2002/03-2003/04	5.75	5.81	-0.06
Average	2.75	2.78	-0.03

5.4 Survival adjustments: hospital output

5.4.1 Simple survival adjustment

This section considers the results of applying the survival adjustment formula in section 4.8.1 to HES data. There are two choices of death rates that can be used in the calculations, those that occur during the hospital stay or in-hospital deaths together with those occurring within some period following discharge from hospital. In-hospital deaths are those most directly attributable to the NHS but are likely to underestimate survival changes due to medical treatment since many patients die within a short time after discharge. However using mortality rates after discharge runs the risk of attributing deaths from extraneous influences to the NHS. On average in the period under consideration 30 day mortality rates were about 25% higher than in-hospital deaths (Table 5.4). Both indicators show a downward trend with similar rates of decline.

Table 5.4 Mortality rates, (deaths/CIPS), 1998/99-2003/04

	In-hospital	30 day
1998/99	0.0239	0.0308
1999/00	0.0238	0.0306
2000/01	0.0229	0.0293
2001/02	0.0236	0.0299
2002/03	0.0228	0.0286
2003/04	0.0222	0.0276

Death rates vary enormously across procedures. Death rates are considerably higher for non-elective procedures than for electives. Although the rate of decline is greater in the latter - on average elective mortality rates declined by 5.9% from 1998/99-2003/04 against decreases of 2.1% for non-electives - aggregate trends are dominated by those for non-electives given their greater weight (Figure 5.3).

Figure 5.3 30 day Mortality rates, electives and non-electives

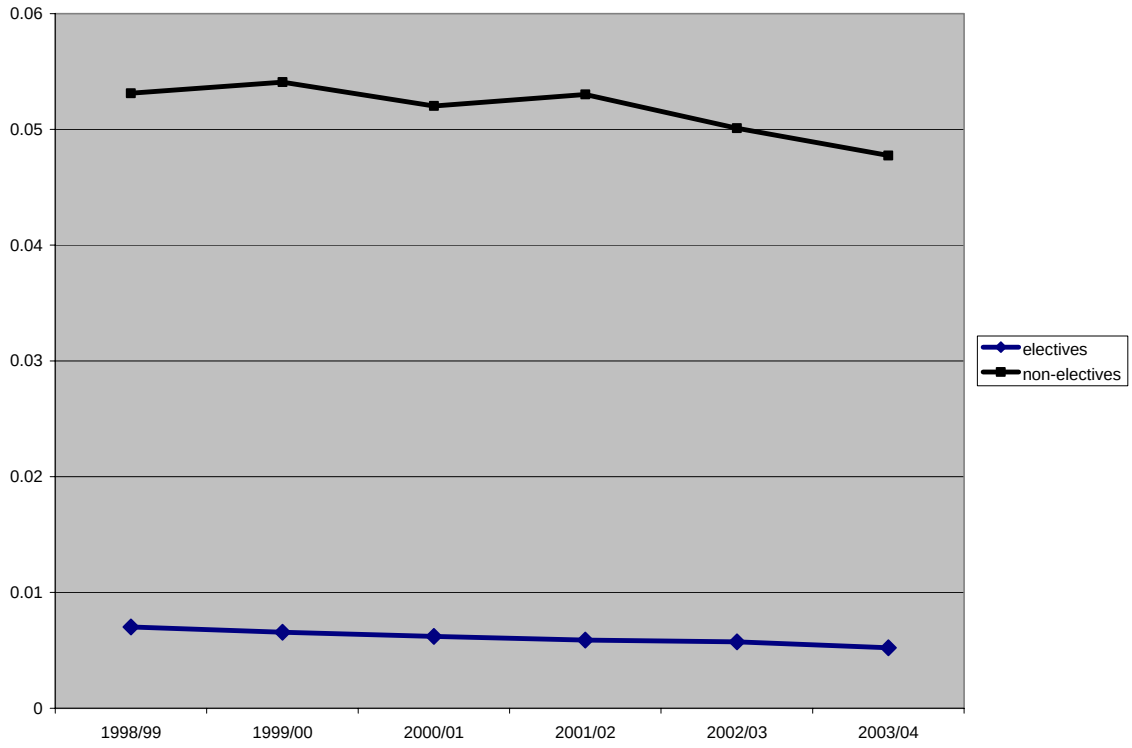


Figure 5.4 Plot of mortality rates 2002/03

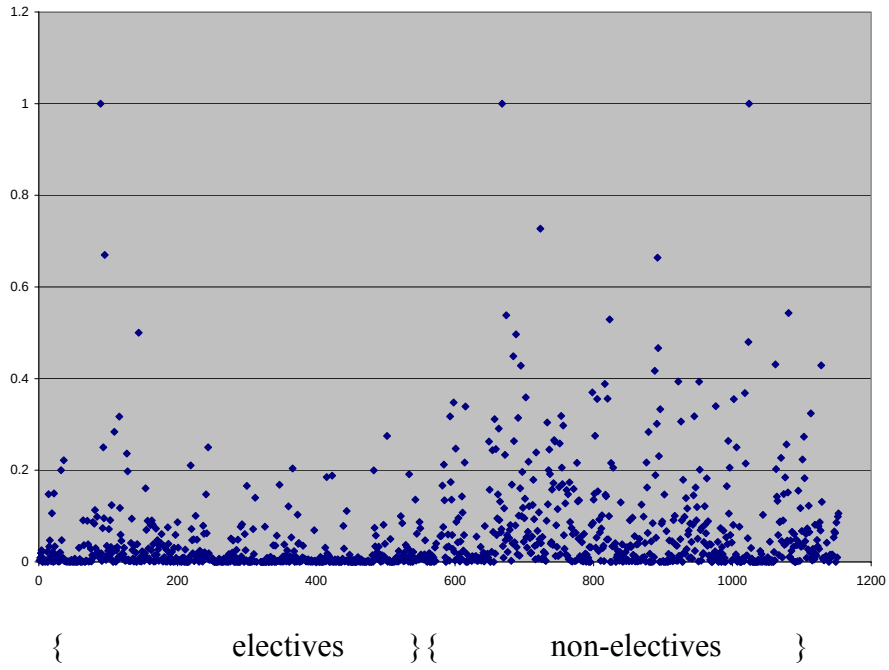


Table 5.5 reports calculations of the pure short term survival adjusted cost weighted output index (section 4.8.1)

$$\frac{\sum_j x_{jt+1}\left(\frac{a_{jt+1}}{a_{jt}}\right)c_{jt}}{\sum_j x_{jt}c_{jt}} \qquad (99)$$

where a is the survival rate.

The first column of results is the unadjusted CWOI. The second and third columns are the survival adjusted indices calculated with 30 day and in-hospital death rates. The adjustments are non-trivial, though generally quite small in percentage point terms. These adjustments are generally larger in the final two years than in the beginning of the period. The use of 30 day mortality rates yields a higher adjustment than in-hospital deaths in all but the first year.

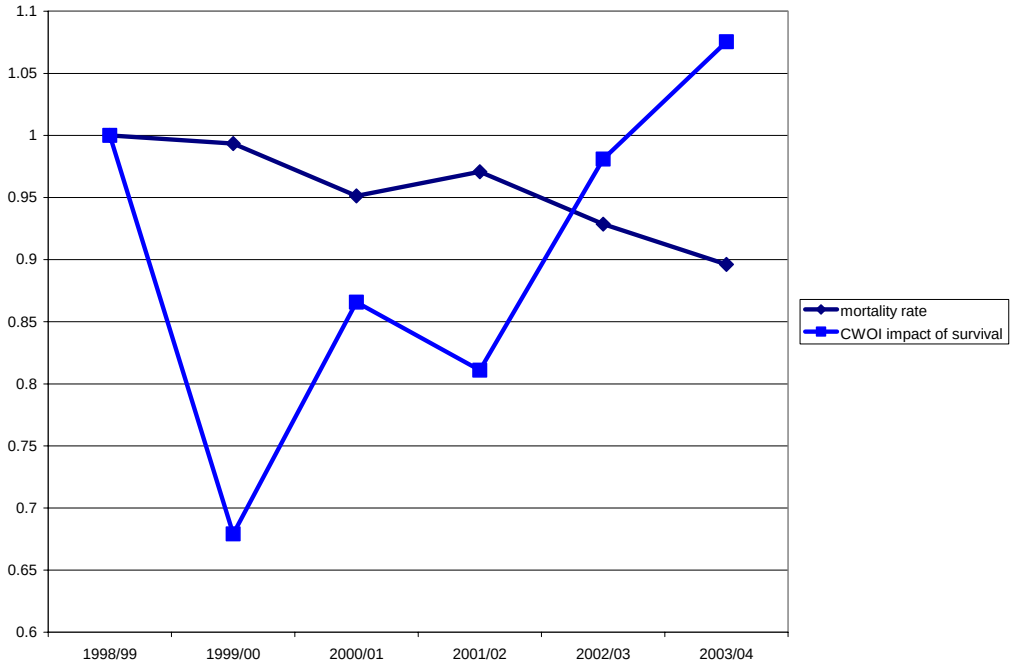
Table 5.5 Laspeyres CWOI index, CIPS, adjusted for survival

	Laspeyres CWOI		
	Unadjusted	Adjusted for survival (1-m)	
		30 day	In-hospital
1998/99-1999/00	1.87	1.27	1.37
1999/00-2000/01	0.91	1.16	1.08
2000/01-2001/02	0.95	0.89	0.86
2001/02-2002/03	4.44	5.37	5.14
2002/03-2003/04	5.81	6.37	6.22
Average all years	2.78	2.99	2.91

The impact of the survival adjustment depends on both the rate of change of survival across HRGs and their cost shares. The latter turn out to have a large impact since, as stated earlier, the majority of procedures show little change in survival but these tend to be concentrated in low cost procedures. To illustrate this point Figure 5.5 shows the change in average (unweighted) mortality rates (from Table 5.4) and the change in the CWOI adjusted for survival minus the unadjusted CWOI (the second column in Table 5.5 minus the first column in Table 5.5), both indexed at 1998 =1. The mortality rate shows a relatively smooth pattern, generally declining but with a small upward shift comparing 2000/01 and 2001/02. In contrast the impact on the CWOI is much more variable, and not always in the inverse direction to the change in the mortality rate.

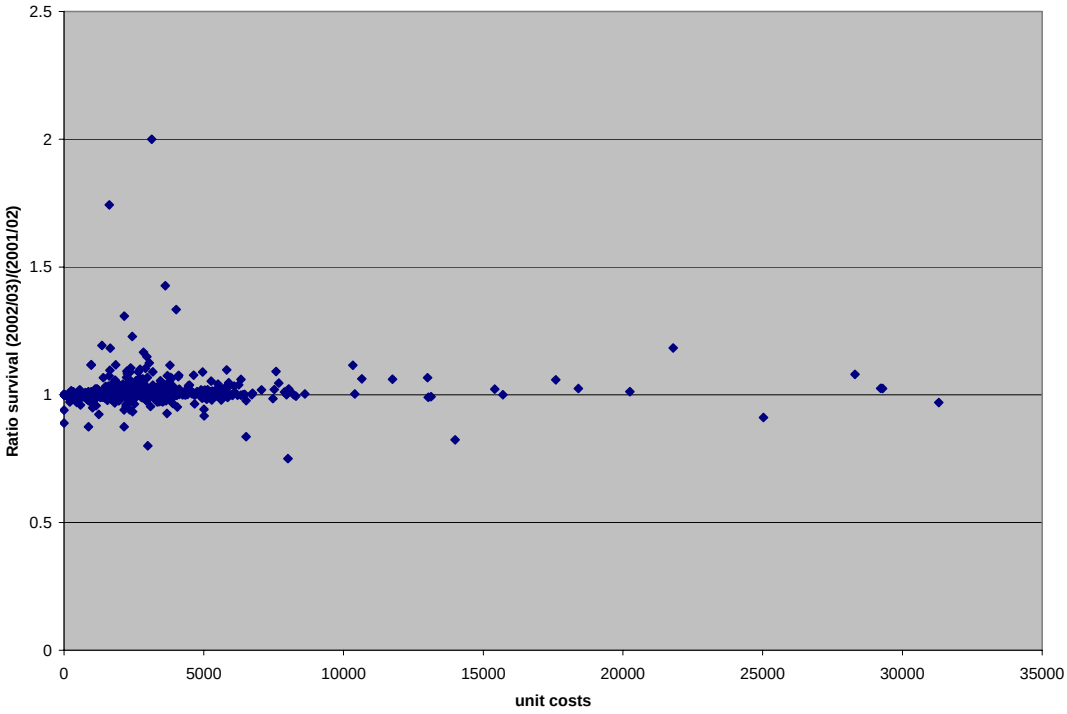
Figure 5.6 plots changes in survival rates against unit cost for one of the growth periods, 2001/02-2002/03. Most changes in survival are small, ranging around the value 1 on the y-axis and the majority of these are in the lowest unit cost range. Year on year changes in the CWOI are driven largely by variations in survival rates in the relatively few procedures with very high unit costs, plus a few cases where changes in survival rates are very high in the low unit cost range.

Figure 5.5 Mortality rates and the impact of the survival adjustment*, Index 1998/99=1



* calculated as the difference between the 30 day survival adjusted CWOI and the unadjusted CWOI

Figure 5.6 Growth in survival (ratio) and unit costs, 2001/02-2002/03, (electives and non-electives)



To understand the sensitivity of the results to cost shares we estimated the change in the index when survival was assumed unchanged for the top 25 high cost share HRGs, which represented just over 30% of total expenditure. The impact of this was to reduce the 30 day survival adjustments by about 60%. Thus the calculations depend heavily on the survival rates of a small number of HRGs. Within this high cost share group, comparing 1999/00 with 1998/99, 17 of the 25 HRGs showed reductions in survival rates and these are responsible to a large extent for the big negative impact of the survival adjustment on the CWOI in that growth period. In contrast in the final two growth periods the majority of high cost share HRGs witnessed increases in survival rates – 19 HRGs in 2001/02-2002/03 and 20 HRGs in 2002/03-2003/04.

Over time, both the number of HRGs with positive growth in survival rates and the share of expenditure accounted by these procedures have increased as shown in Table 5.6. If the percent of HRGs with increases in survival rates is lower than the cumulative expenditure share (in percent) of these procedures, then increased survival is concentrated in relatively high cost procedures. Table 5.6 shows that this is the case in each growth period except the first and that the discrepancy has increased through time.

Table 5.6 Changes in 30 day survival rates and expenditures shares

	Percent of procedures* with change in survival rates >1	Expenditure shares of procedures* with change in survival rates >1
1998/99-1999/00	42.8	37.8
1999/00-2000/01	55.5	62.0
2000/01-2001/02	49.5	50.7
2001/02-2002/03	62.9	75.4
2002/03-2003/04	63.0	77.7

* Total number of procedures = 1148, with electives and non-elective HRGs treated as separate procedures.

We suggested in section 4.8.1 that a simple survival estimate would be likely to be a conservative estimate of the effect of the growth in the health effect of treatment and we now consider how making strong assumptions about the health effect alters the results.

5.4.2 Survival and estimated health effects adjustment

We consider the survival and health effects adjusted index of section 4.8.2

$$\frac{\sum_j x_{jt+1}\left(\frac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)c_{jt}}{\sum_j x_{jt}c_{jt}} \qquad (100)$$

where $k_j = q^o_{jt} / q^*_{jt}$ is an estimate of the proportionate effect of treatment conditional on survival to no treatment which, in the absence data on actual health effects, we assume is constant over time. (q^*_{jt} is the sum of discounted quality adjusted life years accruing to patients who survive treatment. q^o_{jt} is the sum of quality adjusted life years for untreated patients). With $k = 0$ which implies that the patient would have zero quality adjusted life years if not treated we have the pure survival adjusted index. We examine the impact of assuming that k is positive. As noted in section 4.8.2, the rather sketchy available evidence suggests a value of around $k = 0.8$ for non life threatening procedures. When the treatment has a high mortality we set $k = 0$. If we used m = 0.2 for the cut-off mortality rate for setting $k = 0$ this would ensure that term $a - k$ is never negative which would correspond to treatment having a negative effect on health. But as we noted in the simulations in section 4.8.2 this would make the index very sensitive to change in mortality when the rate is close to 0.2. We therefore set the cut off value for mortality which leads to $k = 0$ so that $a - k$ is never smaller than 0.05. Thus for HRGs with high mortality we adjust only by the survival rates, not by the survival rates and the assumed health effect.

Table 5.7 shows that including the crude health effects adjustment via $k = q^o/q^*$ generally increases the growth rate compared with no adjustment (first column) and with a simple survival adjustment (Table 5.5). The greatest impact is the third column, in particular for the final two periods. The results in general suggest that

adjusting for survival adds about two percentage points to the growth rate in 2002/03 and 2003/04. Averaged across the five yearly growth rates, the impact ranges from adding about 1.0 to 0.4 percentage points to the growth rate. Contrast this with an average impact of 0.22 for the simple survival adjustment using 30 day survival rates. Thus a survival adjustment which incorporates crude but not implausible adjustments for health effects is capable of significantly adding to the growth rate of hospital output. Note, as with the simple survival adjustment, much of the impact is due to the behaviour of survival rates in the high cost share HRGs. For example in the case where k=0.8 with cut off = 0.10, nearly 70% of the adjustment can be attributed to the 25 HRGs with the highest cost shares.

Table 5.7 CWOI index, CIPS, adjusted for survival, 30 day mortality rates

	Unadjusted	q^0/q^*=0.8 if m<0.10, q^0/q =0 otherwise	q^0/q^*=0.8 if m<0.15, q^0/q =0 otherwise	q^0/q^*=0.7 if m<0.15, q^0/q =0 otherwise	q^0/q^*=0.7 if m<0.10, q^0/q =0 otherwise	q^0/q^*=0.9 if m<0.05, q^0/q =0 otherwise
1998/99-1999/00	1.87	0.78	0.09	0.73	1.02	1.26
1999/00-2000/01	0.91	1.58	1.97	1.51	1.36	1.54
2000/01-2001/02	0.95	0.91	1.01	0.93	0.90	1.01
2001/02-2002/03	4.44	6.59	7.72	6.34	5.97	6.27
2002/03-2003/04	5.81	7.15	8.04	7.10	6.76	7.09
Average all years	2.78	3.36	3.77	3.28	3.20	3.43

5.4.3 Survival adjustments with health effects and life expectancy

In section 4.8.3 we suggested that including a term reflecting life expectancy of patients treated would be a way of improving the crude adjustment for the health effect and proposed the index

$$\frac{\sum_j x_{jt+1} c_{jt} \frac{\left(a_{jt+1} - k_j\right)}{\left(a_{jt} - k_j\right)} \left(\frac{1 - e^{-rL_{jt+1}}}{1 - e^{-rL_{jt}}} \right)}{\sum_j x_{jt} c_{jt}} \qquad (101)$$

where L_{jt} is the life expectancy at the average age of patients getting treatment j and r is the discount rate on quality adjusted life years (the units in which health effects are measured). Table 5.8 reports the results of calculation of this index with a discount rate on remaining life equal to 1.5% for the simple survival adjustment and with our central case value of $k = 0.8$ with a mortality cut off of either 0.15 or 0.10.

Table 5.8 CWOI index, CIPS, adjusted for survival, life expectancy, 30 day mortality rates, r=1.5

	Unadjusted	q^0/q^*=0.8 if m<0.10, q^0/q =0 otherwise	q^0/q^*=0.8 if m<0.15, q^0/q =0 otherwise
1998/99-1999/00	1.87	1.12	0.74
1999/00-2000/01	0.91	1.37	1.76
2000/01-2001/02	0.95	0.76	0.89
2001/02-2002/03	4.44	6.31	7.44
2002/03-2003/04	5.81	7.13	8.03
Average all years	2.78	3.30	3.72

The growth is higher with either of the adjustments than without them, more markedly for the variant with the more generous cut off which leaves more HRGs being adjusted by the ratio of health effects than by the simple survival ratio. The effect of the life expectancy adjustment is to reduce the growth compared with the corresponding case in Table 5.7 in all years except the first and reflects the increasing age of patients treated by the NHS.

5.5 Waiting time and survival adjustments: hospital output

This section considers additional impacts on the indices from taking account of changes in waiting times. It shows results for a number of variants based on the

formulae in section 4.10. We calculated the mean wait after truncating very long waits to four years and the "certainty equivalent" wait which was the mean plus a "risk premium" to reflect the disutility from the risk of long wait relative to the mean. We used the rule of thumb that the certainty equivalent wait for a treatment was at the 80th percentile wait for that treatment (see section 4.10.4).

Table 5.9 shows mean waits across all patients and the mean 80th percentile wait across HRGs for electives in the period under study. This shows a decline in average waiting times using both measures since 1998/99. However this is mainly due to a large drop between 1998/99 and 1999/2000. Starting in the latter year mean waiting times increased up to 2002/03 but declined marginally in the final year.

Table 5.9 Trends in waiting time, days, averages across HRGs

	Mean		Per cent HRGs with decline in waiting times	
	Truncated mean	80th percentile	Truncated mean	80th percentile
1998/99	88.7	132.2		
1999/00	80.8	117.7	62.3	55.1
2000/01	82.3	119.0	32.4	34.2
2001/02	85.2	124.4	31.9	33.9
2002/03	88.5	128.9	32.3	29.9
2003/04	85.9	126.8	63.4	51.5

The third and fourth column of Table 5.9 summarise the variation across HRGs in terms of waiting time experience by showing the percent of HRGs that show declines in average waiting times. Only in the first and last growth period do the majority of HRGs show decreasing waits with increases in the majority in the intervening years. The overall mean measures of waiting times are affected by the extent to which activity moves between procedures. Although substantial proportions of HRGs record reductions in waits in any one year this does not imply a substantial reduction in waiting times. Indeed if yearly waiting times were symmetrically randomly distributed around an unchanging mean for each HRG then 50% would have

reductions in any given year but there would be no overall downward trend in waiting times. The first two columns of Table 5.9 suggest that adjustments for reductions in waiting times are likely to have a small effect because the measures of waiting time that we use did not change very much.

5.5.1 Effect of waiting time adjustments

The last point above is confirmed when we consider variations in quality adjustments to take account of changes in waiting times within the framework of a cost weighted output index. Tables below show the results of using various scaling factor waiting time formulae from sections 4.10.2, with the two measures of waiting time and different discount rates. The first column of figures in each panel shows as the base case the survival adjustment variant with $k = q^0/q^* = 0.8$ and the mortality cut off set to $m = 0.10$ as a point of comparison. We found that other survival adjustments made little difference to the effects of the waiting time adjustments.

Table 5.10 reports results from the adjustment with discounting to date of treatment with charge for wait (section 4.10.2.2)

$$\frac{\sum_j x_{jt+1} c_{jt} \left(\frac{a_{jt+1} - k_j}{a_{jt} - k_j} \right) \left[\frac{\left(1 - e^{-r_L L_{jt+1}}\right)}{r_L} - \frac{\left(e^{r_w w_{jt+1}} - 1\right)}{r_w} \right]}{\sum_j x_{jt} c_{jt} \left[\frac{\left(1 - e^{-r_L L_{jt}}\right)}{r_L} - \frac{\left(e^{r_w w_{jt}} - 1\right)}{r_w} \right]} \qquad (102)$$

where w_{jt} is the waiting time measure for HRG j, r_w is the discount rate on waiting times and r_L is the discount rate on QALYs. Note this formula differs from that in Table 4.8 due to different discount rates on waits and QALYs – if the two discount rates are equal the formula reduces to that in Table 4.8. The panels differ in the measure of waiting time adopted (mean wait or 80th percentile).

Table 5.10 Laspeyres CWOI index, CIPS, adjustments for changes in waiting times

Discount to date of treatment with charge for wait (based on mean wait variable)					
	Survival adjustment only	$r_w = r_L =$ 1.5%	$r_w = r_L =$ 5%	$r_w =$10%, $r_L =$1.5%	$r_w =$50%, $r_L =$1.5%
1998/99-1999/00	1.12	1.16	1.08	1.16	1.16
1999/00-2000/01	1.37	1.35	1.46	1.35	1.34
2000/01-2001/02	0.76	0.75	0.79	0.75	0.75
2001/02-2002/03	6.31	6.31	6.40	6.32	6.32
2002/03-2003/04	7.13	7.20	7.26	7.20	7.21
Average	3.30	3.32	3.36	3.32	3.32
Discount to date of treatment with charge for wait (based on 80th percentile wait variable)					
	Survival adjustment only	$r_w = r_L =$ 1.5%	$r_w = r_L =$ 5%	$r_w =$10%, $r_L =$1.5%	$r_w =$50%, $r_L =$1.5%
1998/99-1999/00	1.12	1.18	1.13	1.19	1.21
1999/00-2000/01	1.37	1.34	1.44	1.34	1.33
2000/01-2001/02	0.76	0.75	0.78	0.75	0.75
2001/02-2002/03	6.31	6.34	6.44	6.35	6.38
2002/03-2003/04	7.13	7.24	7.31	7.25	7.30
Average	3.30	3.33	3.38	3.34	3.36

Note: All columns have the same survival adjustment: k = 0.8 if m < 0.10, 0 otherwise

We also considered a number of variations in quality adjustments to take account of changes in waiting times based on use of additional formula or different ways of measuring waiting times. The first reports the results for waiting time adjustment with discounting to the date placed on the list (section 4.10.2.1):

$$\frac{\sum_j c_{jt} x_{jt+1} \left(\frac{a_{jt+1} - k}{a_{jt} - k} \right) \left(\frac{e^{-r_w w_{jt+1}} \left(1 - e^{-r_L L_{jt+1}}\right)}{e^{-r_w w_{jt}} \left(1 - e^{-r_L L_{jt}}\right)} \right)}{\sum_j c_{jt} x_{jt}} \qquad (103)$$

where the waiting time adopted is the 80th percentile.

The second is based on the use of individual data to measure mean waiting times. Since the waiting times and life expectancy factors are non-linear and there is a variation in waiting times and in ages within an HRG in a given year it is possible that our use of a single waiting time and life expectancy estimate for each HRG may lead to misleading results. We therefore computed the equivalent of the waiting time adjustment with discounting to date of treatment with a charge for waiting with individual level data.

Thirdly we were asked to consider how an adjustment for waiting times could allow for optimal waiting times – it was suggested that some patients might find too short a wait inconvenient. In the absence of any information on what an optimal wait might be we investigated the implications of assuming that the effect of an optimal waiting time w* was to replace the actual wait in our waiting time adjustments with the $\hat{w} = w - w^*$ if $w > w^*$ and 0 otherwise. Thus reductions in waiting time below w^* would have no effect whereas the proportionate effect of reductions above w^* would be increased. We experimented first with $w^* = 30$ days but found that this resulted in a large number of HRGs where $\hat{w} = 0$, therefore we opted to use a value of w = 15 days.

Table 5.11 shows the impact on the CWOI of these three variants, where the first column shows the calculations in Table 5.10, discount to date of treatment with charge for wait (based on 80th percentile wait variable) for comparable discount rates. Discounting to date on list lowers the average growth rates, mainly through reductions in the first and last years. The use of individual data has a greater effect in raising the growth rate, although this is concentrated in the first few years. The use of optimal waits has little impact on the average growth rates, with only a discernible impact in the final year.

Table 5.11 Laspeyres CWOI index, CIPS, adjustments for changes in waiting times, $r_L = r_W = 1.5\%$

	Discount to date of treatment with charge for wait, 80th percentile wait variable	Discount to date on list, 80th percentile wait	Using individual data, discounting to date of treatment with charge for wait	Optimal waits* discounting to date of treatment with charge for wait, mean wait
1998/99-1999/00	1.18	0.69	1.30	1.18
1999/00-2000/01	1.34	1.36	1.59	1.34
2000/01-2001/02	0.75	0.74	1.05	0.75
2001/02-2002/03	6.34	6.53	6.41	6.34
2002/03-2003/04	7.24	7.07	7.24	7.25
Average	3.33	3.24	3.48	3.33

*Based on 15 day optimal waiting time.
Note: All columns have the same survival adjustment: k = 0.8 if m < 0.10, 0 otherwise

The results which show small effects of waiting time adjustments are largely driven by the lack of change in waiting times rather than the methods used. To see this suppose waiting times for the 80th percentile were reduced by 10% for all HRGs comparing 2003/04 with 2002/03. Then the discount to date of treatment with low discount rates equal to 1.5% would add 0.16 percentage points. With the same discount rates, reducing waits at the 80th percentile by 50% would add 1.12 percentage points. The results from the specimen index calculated with a much smaller set of HRGs are sensitive to the method of waiting time adjustment and can make a difference to estimated growth rates.

In addition the impact of changes in waiting times is dependent on the cost share weights. In this case however, large increases in waiting times tend to be concentrated in low unit cost procedures. This is illustrated the final two growth periods in Figure 5.7 and Figure 5.8 but a similar pattern is also apparent for earlier years.

Figure 5.7 Percentage changes in waiting times (days) and unit costs, 2001/02-2002/03

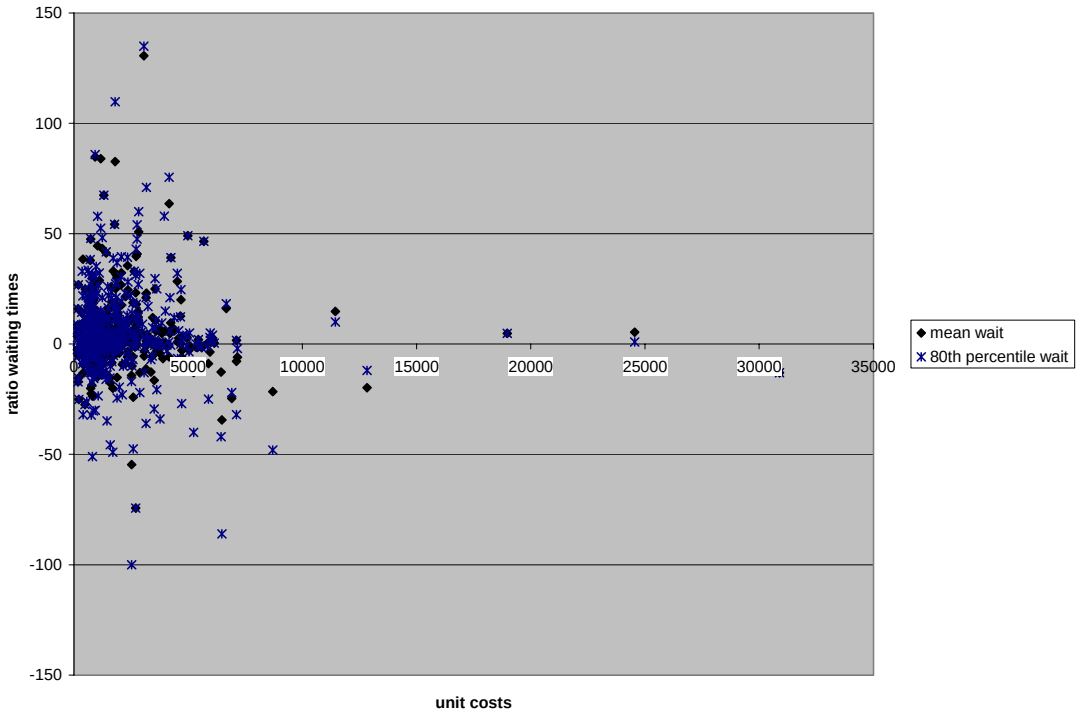
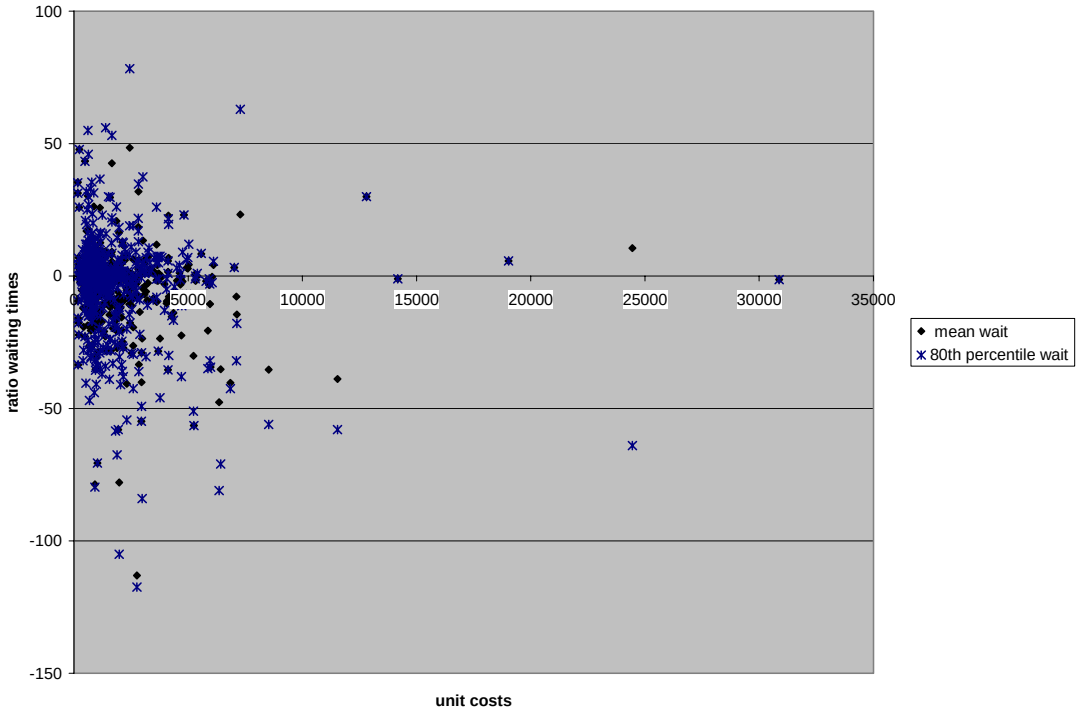


Figure 5.8 Percentage changes in waiting times (days) and unit costs, 2002/03-2003/04



Again it is useful to summarise the relationship between cost and changes in waiting times by the number of HRGs that show reductions and their expenditure shares. Table 5.12 shows that for three of the five growth periods the majority of HRGs show increases in waiting times with higher proportions in the first and final period. In general the percent of HRGs with reductions in waiting times are about equal to their expenditure shares so that reductions tend to be concentrated at the low unit cost end.

Table 5.12 Changes in waiting times and expenditures shares

	Mean wait		80th percentile wait	
	Per cent of electives* with reduction in waiting time	Expenditure shares of electives* with reduction in waiting time	Per cent of electives* with reduction in waiting time	Expenditure shares of electives* with reduction in waiting time
1998/99-1999/00	62.3	68.7	55.1	65.8
1999/00-2000/01	32.4	37.3	34.2	37.7
2000/01-2001/02	31.9	30.4	33.9	35.2
2001/02-2002/03	32.3	33.3	29.9	29.8
2002/03-2003/04	63.4	63.9	51.5	51.7

* Total number of electives = 563

We did not estimate the alternative characteristic adjustment set out in section 4.10.1 for all HRGs because of lack of data on health effects. However, we report in section 6 results from using this approach to waiting times with a small specimen set of HRGs for which we have better health data.

5.5.2 Outpatient waits

Finally we consider outpatient waits. Data on waiting times for first outpatient attendances are only available for four of the years considered in this report. Average days wait for outpatients were 64 days in 1999/00 and 2000/01 but then declined by about 10% in 2002/03 to 58 days and a further 7% to 54 days in 2003/04. We used the discount to date of treatment formula as for electives above, assuming all outpatients had remaining life expectancy of 26 years, the average across electives. The cost weights for changes in waiting times for outpatients was assumed to be the

sum of the cost share of first attenders and follow up appointments to be consistent with the spells approach employed in previous calculations. The effect of this adjustment was to increase the cost weighted output index for outpatient first attenders from 4.47% to 4.59% in 2001/02 and from 6.48% to 6.56% in 2002/03. These adjustments become very small when all outpatients including follow-ups are included in the index.

5.6 Additional quality adjustments

We also considered the use of data in addition to survival and waiting time in order to quality adjust the output index. These additional adjustments are necessarily speculative because of the absence of crucial data so we present them mainly to illustrate the application of the methods described in section 4 and to give a very rough indication of what are the crucial parameters on which information is required Given current data availability we do not recommend they be used to quality adjust the NHS output index. The adjustments are of two types (a) we treat measures of readmissions and MRSA as indicators of unnecessary additional expenditure (section 4.9.1); and (b) we use measures of patient experience as summary indicators of characteristics that patients value (section 4.11) .

5.6.1 Adjusting for the costs of poor treatment: readmissions and MRSA

We suggested in section 4.9.1 that one way of accounting for readmissions and MRSA was to argue that these led to lost output whose value, in accordance with the assumptions underlying the cost weighted index, was their additional cost to the NHS. Thus, ignoring other quality adjustments for illustrative purposes, we calculate

$$\frac{\sum_j x_{jt+1}c_{jt} - \sum_j x^b_{jt+1}c^b_{jt}}{\sum_j x_{jt}c_{jt} - \sum_j x^b_{jt}c^b_{jt}} \qquad (104)$$

where x^b denotes the number of readmissions or cases of MRSA and c^b their costs. As we noted in our discussion in section 4.9.1 (see also Appendix A) the current data on the x^b are not sufficiently detailed, to enable us to distinguish say between readmissions which are the result of poor initial treatment and those which result from pre-existing poor health of the patient. In our calculation we therefore use the total

number of readmissions and the total number of MRSA cases. Since we have no data on MRSA cases or readmissions prior to 2001/02 we show only the effect from 2001/02 to 2003/04.

There are also problems in estimating the costs c^b if we do not know, for example, which readmissions are indicators of poor treatment. We therefore use notional costs of a readmission of £500 and of £1000 for an MRSA case at 2002/03 prices, with prices in other years estimated using money GDP per capita.

With these data and working from the cost weighted index with no adjustment for mortality or waiting time, we can see what effect these have on the estimated growth rate. We show three cases in addition to the basic cost weighted index.

Table 5.13 Effects of quality adjustment for readmissions and MRSA on cost weighted hospital output

	No. of Readmissions	MRSA Cases	Hospital CWOI*	Adjusted Indices		
MRSA Charge (£ per case)			0	£1,000	£0	£1,000
Readmission Charge (£ per case)			0	£0	£500	£500
2001/2	476,556	17,933				
2002/3	492,247	18,519	4.44%	4.44%	4.46%	4.46%
2003/4	536,005	19,311	5.81%	5.81%	5.75%	5.75%
Average growth			5.12%	5.12%	5.10%	5.11%

*This is the unadjusted CWOI, for the inpatient hospital sector (see Table 5.3)

The impact of MRSA cases is negligible. The number of cases is very small, compared to the number of patients treated in hospitals; a very much higher cost would be required for it to have an impact. The effect of readmissions is slightly larger (see section 4.9). However since the growth rate, at 6% p.a. over the two years is not very different from the growth rate of the unadjusted impact, the costs associated with readmissions (which are regarded as money wasted rather than contributing to output) would have to be very large for there to be a substantial effect on the overall index.

Table 5.14 shows the effect of adding the readmission and MRSA adjustments to the survival and waiting time adjusted hospital cost weighted output index. It also further clarifies the calculations, in the case where both adjustments are included, by showing the shares of total expenditure attributed to each. With such small shares, especially for MRSA, it is unlikely that the impact on the overall index would ever be of great significance.

Table 5.14 Effects of quality adjustment for readmissions and MRSA on cost weighted inpatient hospital output in addition to survival and waiting time adjustment

	Index adjusted for Waiting and Mortality	Index Additionally Adjusted for Readmissions and MRSA	Growth Rate MRSA	Growth Rate Readmissions	Assumed Costs of MRSA as Proportion of Total Costs	Assumed Costs of Readmissions as Proportion of Total Costs
2001/02					0.14%	1.89%
2002/03	6.35%	6.41%	3.27%	3.29%	0.14%	1.81%
2003/04	7.25%	7.22%	4.28%	8.89%	Combined Growth Rate MRSA/Readmission	
Av 03/4 over 01/02	6.80%	6.82%	3.77%	6.05%	5.90%	

Note: The index adjusted for waiting and mortality is the 80th percentile waiting variable with r_w=10% p.a. and r_l=1.5% p.a. in table 5.10. The adjustments for readmissions and MRSA are incorporated in the index by adding to the growth in the index adjusted only for waiting and mortality the growth rates of MRSA and readmission weighted negatively by the cost shares for the previous year.

These speculative guesstimates suggest that adjusting for readmissions and MRSA in this way can have non-trivial effect on the survival and waiting time adjusted hospital cost weighted output index. While we stress the illustrative nature of these figures one important policy point does follow from them. Cases of MRSA are rare that costs associated with its treatment would have to be a substantial multiple of the £1000 we assumed before it could have an important impact on the index. Readmission, on the other hand, is of material importance. If the DH wishes to use our approach to adjust the CWOI we recommend that it should focus initially on quantifying the costs of MRSA cases and the proportion of readmissions which are avoidable or harmful.

5.6.2 Patient satisfaction

As we note in our discussion in section 4.11, there are theoretical arguments against using such data, not least that it may simply be double counting aspects of quality, such as health effects and waiting times, which can be captured by other more direct means. On the other hand if one believes that satisfaction survey response measure characteristics of NHS care which are of value to patients and are not already reflected in other quality measures then we have suggested in section 4.11 a method of incorporating such data. The data are described in Appendix A.

Since patient satisfaction data only permit comparison of 2004 or 2004/05 with 2003 we have illustrated our method by examining its impact on the average annual growth rate in hospital output between 2001/02 to 2003/04.[8]

Our method has three main steps. First we quantify the ordered qualitative responses to various satisfaction questions by assigning them equally spaced numerical values between 0 and 100 for the least to the most satisfied categories. We construct such numerical scores for three aspects of the patient experience: food, cleanliness and non-clinical experience (for example whether patients felt they were treated with respect and dignity). We have scores from A&E, outpatient, and inpatients surveys. There is a residual category of "other" hospital activity, taking up about 25% of costs. We have assumed that the surveys do not describe patient satisfaction with the services provided by this.

Second, since these scores are in effect estimates on a per patient basis we need to scale them by a suitable measure of the volume of such experiences. We use the number of patients for outpatients and A&E and the number of patient spells for inpatients. The reason for choosing patient spells for inpatients rather than the alternative of bed-days is that the analogy with hotel services can be taken only so far. One can argue that most patients would prefer short stays rather than long stays in hospital- that staying in hospital is a necessary evil rather than a hotel service consumed with the readiness of a stay in a hotel. For practical purposes, because we

[8] We also calculated patient satisfaction adjustments to the volume of patient consultations for the same period which had little effect because the patient experience scores changed little over the period

feel that, because a stay in hospital is a route to better health rather than a consumption good *per se*, we make the calculation with the volume of treatment measured by numbers of consultant inpatient spells.

Third, the scores are incorporated into the output index in two ways. We use expenditure on food and cleaning to weight the food and cleanliness scores, taking account of the shares of A&E, outpatients, and inpatients in total hospital costs in 2001/02.[9] We cannot, however find from the expenditure data weights appropriate to non-clinical experience. We therefore present results making the assumption that either 5% (case A) or 10% (case B) of total expenditure by hospital trusts and that these same proportions apply to the three activities covered. It can be doubted whether any form of accounting would identify the proportions since the score includes measures of politeness and courtesy which cost nothing.

Table 5.15 reports an illustrative calculation of growth rates in the satisfaction scores, and the growth in the total volume of quality taking account of the numbers experiencing these different aspects of care in the different sectors.

[9] These weights overstate the importance of the cost of providing the "hotel service" components of cleanliness and food quality since they also have medical consequences. But since we do not have data on the consequences of medical treatment for quality of life, we are not in fact double-counting. In any case, since the cost weights are small, double-counting is unlikely to be a major source of error.

Table 5.15 Illustrative indicators of patient satisfaction with hospital services, average annual growth rates, 2001/02 to 2003/04

	A&E	Outpatients	Inpatients	Overall Change	Weight in overall Trust Budget	
					A	B
Food			0.64%	0.64%	0.85%	0.85%
Cleanliness	-1.02%	-2.27%	-0.63%	-1.05%	1.44%	1.44%
Superficial Attention	0.70%	-0.10%	0.56%	0.40%	5.00%	10.00%
Quality Change (A)				0.14%		
Quality Change (B)				0.25%		
Volume Change	4.59%	5.05%	5.12%	5.08%		
Total Change (A)				5.23%	7.29%	
Total Change (B)				5.49%		12.29%
Weight	4.47%	24.21%	71.32%			

Notes:

1. The changes in the indicators of food, cleanliness and non-clinical care are the changes in the logarithms of the relevant variables. This is also true of the changes shown in tables A1 to A4.
2. The changes in each indicator for each category of treatment are weighted together using the weights in the last row so as to give the overall change in each quality attribute. The food indicator is used as it stands because all expenditure on food is associated with inpatients.
3. The overall quality indicators for each attributed are then weighted using weights (A) or (B) shown in the last two columns. To give the overall quality changes (A) and (B). These growth rates are then combined with the overall volume change (calculated as the weighted sum of the changes in the individual volumes) to give the total change using weights (A) and (B).
4. The quality changes for food, cleanliness and non-clinical care are calculated for whatever period the data happen to be available. The changes in volume relate to 2003/04 over 2001/02.

The overall impact on the overall hospital output index of these adjustments is shown in Table 5.16. The effect of the quality terms is further damped because total hospital output includes the "other" activities in addition to inpatients, outpatients and accident and emergency. We have no quality data on these. But in any case, with the MRSA/Readmissions and quality indices both growing at rates not very different from the overall quality-adjusted CWOI, it is not surprising that these effects have little influence on the overall total.

Table 5.16 Illustrative calculations of hospital CWOI with adjustments for survival, waiting times, patient satisfaction as measured in patient surveys, readmissions and MRSA. Average annual growth rates 2001/02 to 2003/04

Average Growth Rates 2001/2 to 2003/4	% p.a.
Unadjusted CWOI	4.34%
Quality Variant 1	5.74%
With Adjustment for Patient Satisfaction (5% weight on non-clinical care satisfaction)	5.71%
With adjustment for MRSA, Readmissions and Patient Satisfaction (5% weight on non-clinical care satisfaction)	5.71%
With Adjustment for Patient Satisfaction (10% weight on non-clinical care satisfaction)	5.69%
With adjustment for MRSA, Readmissions and Patient Quality and Satisfaction (10% weight on non-clinical care satisfaction)	5.69%

Note: these figures refer to the broader definition of the hospital sector; see Table 9.2 and preceding discussion.

5.7 Conclusions

This section has implemented a number of the quality adjustments to hospital sector output based on the methods developed in section 4 to make use of currently available data. We found that

- the pure survival adjustment raises the average annual growth rate of the hospital sector between 1999/00 and 2003/04 from 2.78% to 2.99% when survival was measured at 30 days.
- the survival effect has a smaller effect when calculated using in-hospital survival (average growth 2.91%).
- combining the survival adjustment with an assumed uniform proportional health effect further increase the average growth rate by around 0.4% to 1.0% depending on the assumed value of health effect and the mortality rate cut off criteria used to reduce the volatility of the index.
- adding a life expectancy adjustment to the survival and health effect adjustment had little additional effect on the growth rate.
- combining survival, assumed health effect, waiting time and life expectancy adjustments produced estimated growth rates which were very similar to estimates with only survival and assumed health effect adjustments.
- the effect of the waiting time adjustment was insensitive to very large

variations in discount rates on waiting times, to the use of individual rather than HRG level data, to the form of the adjustment, and to the measure of waiting time (mean wait or 80th percentile wait).

- the small waiting time effects are due to the small changes in waiting times over the period rather than to the form of the waiting time adjustment and the particular parameter values used.
- a crude illustrative adjustment readmissions and MRSA (all that is possible with current data) in addition to the survival, assumed health effect, life expectancy and waiting times adjustments, had no perceptible effect on the average annual growth rate (2001/2 to 2003/4).
- a similarly crude illustrative adjustment for patient satisfaction with food, cleanliness and non-clinical care reduced the growth rate very slightly, (2001/2 to 2003/4) by less than 0.1 percentage point.

6 Specimen output index

6.1 Introduction

In Section 2.4 we set out our preferred index of NHS output, a value weighted output index. It is not possible to estimate a comprehensive value weighted index because of a lack of data on the most important characteristic: improvement in health for patients who survive treatment. In Sections 2.7 and 4.6 we examined the stringent assumptions necessary to justify using cost weights in the output index. In Section 4.8.2 we looked at the sensitivity of the output index to an estimate of health gain based on the assumption that health gain was constant across all activities and did not vary over time.

In this section we examine the implications of having better health data which permit the calculation of our preferred value weighted output index. We have identified a few HRGs where data exist on health outcomes (Appendix C). The data are similar to what would be produced by sampling patients before and after treatment and are used

to drop the restrictive assumption that health gain is constant across activities. While the conditions for which we have outcomes data are not representative of all NHS activities, we are able to compare a number of "specimen" indices with the equivalent cost weighted output index for the same sub-set of conditions.

We use the health data to examine:

- A value weighted output index assigning monetary weights to improvements in health and reductions in waiting times
- The impact on an index of substituting cost weights with value weights
- The effect on health effect adjusted cost weighted indices of allowing health gain to vary by treatment

In the next subsection we describe the data used in the construction of the specimen index. We then estimate the following indices:

- A Cost Weighted Output Index (CWOI)
- CWOI with a short-term survival adjustment
- CWOI incorporating health adjustment
- CWOI incorporating health and waiting times adjustment
- A health outcome weighted output index (HOWOI)
- A HOWOI incorporating waiting times adjustment
- A Value Weighted Output Index (VWOI) where health and waiting times are treated as characteristics

We also explore the sensitivity of results to

- In-hospital versus 30-day survival rates
- The measurement of waiting times
- Discount rate
- Monetary value of a QALY
- Monetary value applied to a day spent waiting

6.2 Data

The specimen index comprises the HRGs listed in table 6.1 below. For all of these HRGs, data on the health outcomes before and after treatment were available, either from clinical trials that employed the EQ5D or our analysis of SF36 from BUPA and York District Trust. The data derived from these two instruments are converted to a common scale. These data are described in greater detail in Appendix C.

Table 6.1 Before and after health outcomes

			Health outcome	
HRG description	**Source**	**HRG**	h_j^0	h_j^*
Intermediate Pain Procedures	BUPA	A07	0.41	0.57
Phakoemulsification Cataract Extraction with Lens Implant	BUPA	B02	0.73	0.76
Other Cataract Extraction with Lens Implant	BUPA	B03	0.70	0.72
Mouth or Throat Procedures - Category 2	BUPA	C14	0.87	0.95
Nose Procedures - Category 3	BUPA	C22	0.83	0.91
Mouth or Throat Procedures - Category 3	BUPA	C24	0.77	0.93
Coronary Bypass	BUPA	E04	0.50	0.73
Acute Myocardial Infarction w/o cc	EQ5D	E12	0.68	0.72
Percutaneous Transluminal Coronary Angioplasty (PTCA)	BUPA	E15	0.54	0.79
Chest Pain >69 or w cc	EQ5D	E35	0.63	0.69
Inguinal Umbilical or Femoral Hernia Repairs >69 or w cc	BUPA	F73	0.64	0.69
Inguinal Umbilical or Femoral Hernia Repairs <70 w/o cc	BUPA	F74	0.74	0.81
Liver Transplant	EQ5D	G01	0.53	0.59
Biliary Tract - Major Procedures >69 or w cc	BUPA	G13	0.63	0.66
Biliary Tract - Major Procedures <70 w/o cc	BUPA	G14	0.68	0.81
Primary Hip Replacement	BUPA	H02	0.37	0.62
Primary Knee Replacement	York	H04	0.35	0.54
Soft Tissue Disorders >69 or w cc	BUPA	H23	0.77	0.84
Soft Tissue Disorders <70 w/o cc	BUPA	H24	0.72	0.74
Inflammatory Spine, Joint or Connective Tissue Disorders <70 w/o cc	EQ5D	H26	0.41	0.53
Complex Breast Reconstruction using Flaps	BUPA	J01	0.93	0.96
Non-Malignant Prostate Disorders	EQ5D	L32	0.81	0.85
Upper Genital Tract Major Procedures	BUPA	M07	0.70	0.80
Threatened or Spontaneous Abortion	BUPA	M09	0.72	0.83
Psychiatric Disorders	EQ5D	P18	0.36	0.41
Varicose Vein Procedures	EQ5D	Q11	0.77	1
Surgery for Degenerative Spinal Disorders	BUPA	R02	0.37	0.67
Spinal Fusion or Decompression Excluding Trauma	BUPA	R03	0.36	0.62
Revisional Spinal Procedures	BUPA	R09	0.32	0.60

Annual data from 1998/99 to 2003/04 on the following are used in the construction of the indices:

- Activity, measured as continuous inpatient provider spells (CIPS), derived from HES.
- In-hospital and 30-day survival rates, derived from HES.
- Waiting times, measured as the mean waiting time and the wait at the 80th percentile, derived from HES.
- Life expectancy, derived from life tables and estimated according to the average age of those in each HRG.

Raw data for each of the variables used in the indices are provided for each year from 1998/99 to 2003/04. To give an intuitive sense of the change in these data over time, and hence what the various indices will be capturing, some of the data are presented in the following tables and figures.

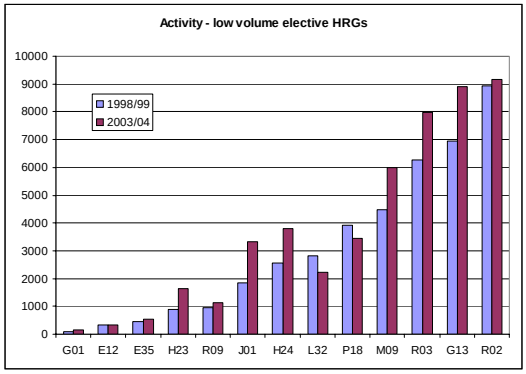
Table 6.2 provides elective and non-elective activity, measured as CIPS, for each year. Figure 6.1 shows the amount of activity in each HRG in 1998/99, when the series begins, and 2003/04, when the series ends. Figures (a) and (b) show the amount of elective CIPS for, respectively, low and high volume HRGs. Figures (c) and (d) provide similar information for non-elective CIPS. For the majority of HRGs, the number of CIPS in 2003/04 (the darker bars) is greater than the number in 1998/99. All else equal, this would be expected to translate into a positive change in the index over the full period.

Table 6.2 Activity by year

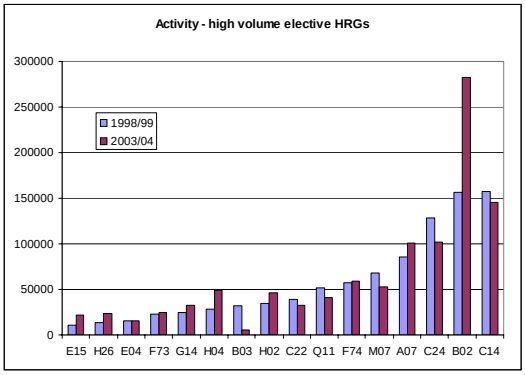
	Elective Activity						Non-elective Activity					
HRG	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**
A07	85,645	86,770	88,495	93,227	99,493	100,640	1,444	1,048	780	679	679	836
B02	155,929	177,551	212,176	224,247	250,377	282,486	386	392	458	574	608	583
B03	32,040	20,644	13,542	10,648	7,755	5,602	172	110	101	66	34	42
C14	157,225	152,462	143,192	140,591	144,338	145,078	6,648	6,665	6,523	6,134	6,395	6,557
C22	39,451	37,603	37,511	31,858	33,653	33,026	7,639	7,680	6,899	6,773	6,919	6,542
C24	128,004	114,666	101,809	100,901	106,716	101,381	6,652	7,006	7,003	6,936	7,121	7,622
E04	15,215	14,618	14,860	15,046	16,280	15,132	2,267	1,141	1,133	950	1,935	2,074
E12	330	212	228	187	328	332	67,422	58,969	57,130	55,455	63,691	63,900
E15	10,555	11,364	13,178	15,439	17,625	21,490	4,672	4,918	5,066	4,995	8,327	10,488
E35	446	412	506	482	534	530	32,875	34,483	40,355	42,724	47,025	51,843
F73	22,300	21,432	21,886	21,348	22,918	24,405	3,322	3,189	3,105	2,957	3,097	3,217
F74	57,034	54,838	55,927	54,517	58,334	59,413	2,995	2,884	2,847	2,782	2,864	3,070
G01	75	99	57	88	112	141	324	386	302	327	343	298
G13	6,935	6,825	7,439	7,600	8,318	8,910	1,406	1,417	1,497	1,534	1,711	1,824
G14	24,491	24,707	26,169	27,636	31,026	33,125	1,879	1,962	2,196	2,264	2,822	3,157
H02	34,122	34,355	36,100	37,530	41,630	46,126	1,262	1,168	1,154	1,149	1,297	1,157
H04	27,741	28,730	31,685	34,392	41,037	48,916	166	159	182	218	205	255
H23	896	1,153	1,364	1,189	1,487	1,637	8,537	8,876	10,066	10,562	11,793	12,911
H24	2,547	3,121	3,523	3,014	3,330	3,807	11,156	11,877	12,887	13,136	13,740	14,527
H26	14,060	14,281	13,495	15,591	21,146	23,992	7,871	7,735	7,435	7,214	7,980	8,076
J01	1,833	2,020	2,481	2,655	3,095	3,337	21	13	30	22	31	14
L32	2,820	2,603	2,497	2,184	2,382	2,219	2,881	2,975	2,652	2,692	3,032	3,355
M07	67,938	62,433	58,150	55,138	53,764	52,315	11,069	10,239	10,225	9,736	9,742	9,889
M09	4,467	4,864	4,913	5,039	5,629	5,998	53,023	55,123	56,158	59,124	63,393	64,311
P18	3,919	3,378	2,862	3,308	3,816	3,439	219	206	231	192	136	192
Q11	51,872	45,659	43,145	40,306	43,846	41,156	158	150	161	150	104	128
R02	8,921	8,368	8,394	8,150	8,808	9,161	1,867	1,589	1,648	1,457	1,608	1,663
R03	6,249	6,029	6,032	6,329	7,015	7,987	1,022	870	874	792	901	1,044
R09	951	924	940	931	1,142	1,135	163	168	158	133	182	195
Average	**33,242**	**32,487**	**32,847**	**33,089**	**35,722**	**37,342**	**8,259**	**8,048**	**8,250**	**8,335**	**9,232**	**9,647**

Figure 6.1 Number of elective and non-elective CIPS by HRG, 1998/99 and 2003/04

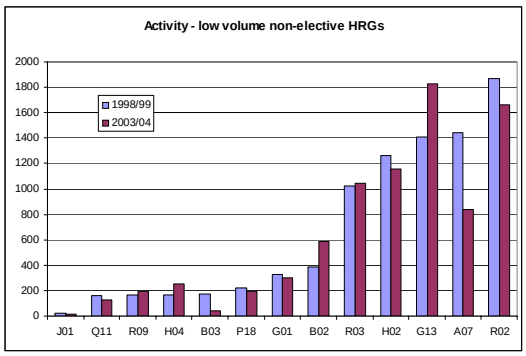
(a) Low volume elective HRGs



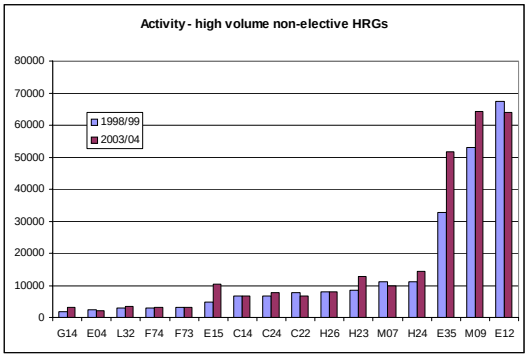
(b) High volume elective HRGs



(c) Low volume non-elective HRGs



(d) High volume non-elective HRGs



There are two available measures of the rates of survival for each HRG – in-hospital survival and 30-day survival. Data on survival are provided in Table 6.3. The average survival rate among electives was upwards of 99% and around 96% for non-electives.

Table 6.3 30-day and in-hospital survival rates, by year

	Elective survival rate											
	30 days						inhospital					
HRG	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**
A07	99.83%	99.87%	99.85%	99.84%	99.87%	99.89%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
B02	99.51%	99.57%	99.62%	99.67%	99.66%	99.69%	100.00%	99.99%	100.00%	100.00%	100.00%	100.00%
B03	99.52%	99.56%	99.59%	99.63%	99.68%	99.77%	99.99%	100.00%	99.97%	100.00%	100.00%	100.00%
C14	99.94%	99.95%	99.94%	99.93%	99.94%	99.95%	99.99%	99.99%	100.00%	99.99%	100.00%	100.00%
C22	99.93%	99.92%	99.94%	99.95%	99.92%	99.94%	99.99%	99.99%	99.99%	99.99%	99.99%	100.00%
C24	99.88%	99.87%	99.88%	99.90%	99.90%	99.91%	99.97%	99.96%	99.96%	99.97%	99.97%	99.97%
E04	98.07%	98.18%	98.27%	98.21%	98.19%	98.66%	98.28%	98.41%	98.52%	98.53%	98.44%	98.86%
E12	72.49%	74.59%	75.90%	72.68%	79.67%	87.42%	73.96%	77.70%	77.19%	75.26%	82.17%	88.23%
E15	99.52%	99.69%	99.59%	99.63%	99.67%	99.65%	99.77%	99.83%	99.79%	99.82%	99.82%	99.84%
E35	97.32%	98.04%	97.46%	98.96%	97.60%	98.91%	99.33%	99.26%	98.44%	99.38%	99.08%	99.27%
F73	99.47%	99.53%	99.55%	99.56%	99.61%	99.63%	99.87%	99.87%	99.88%	99.86%	99.88%	99.90%
F74	99.92%	99.95%	99.95%	99.93%	99.94%	99.95%	100.00%	99.99%	100.00%	99.99%	100.00%	100.00%
G01	94.52%	93.00%	84.75%	90.59%	92.79%	94.24%	94.52%	93.00%	84.75%	90.59%	92.79%	94.24%
G13	98.77%	99.13%	98.73%	99.13%	99.15%	99.20%	99.09%	99.37%	99.01%	99.26%	99.41%	99.37%
G14	99.94%	99.93%	99.92%	99.94%	99.94%	99.95%	99.98%	99.98%	99.98%	99.98%	99.98%	99.99%
H02	99.22%	99.21%	99.27%	99.31%	99.43%	99.33%	99.55%	99.54%	99.53%	99.54%	99.66%	99.59%
H04	99.30%	99.34%	99.26%	99.41%	99.36%	99.50%	99.61%	99.64%	99.56%	99.62%	99.65%	99.67%
H23	97.99%	98.78%	99.34%	99.58%	99.20%	99.33%	99.33%	99.39%	99.49%	99.92%	99.66%	99.70%
H24	99.96%	99.94%	99.91%	99.83%	99.88%	99.82%	99.96%	100.00%	100.00%	99.93%	99.97%	99.95%
H26	99.87%	99.88%	99.90%	99.94%	99.95%	99.93%	99.94%	99.96%	99.94%	99.99%	99.98%	99.98%
J01	100.00%	99.85%	99.96%	99.85%	99.94%	99.94%	100.00%	99.90%	100.00%	99.92%	99.94%	99.97%
L32	99.65%	99.50%	99.48%	99.50%	99.58%	99.73%	99.89%	99.81%	99.84%	99.91%	99.83%	99.91%
M07	99.82%	99.83%	99.81%	99.79%	99.81%	99.82%	99.90%	99.89%	99.87%	99.86%	99.89%	99.88%
M09	100.00%	99.98%	99.96%	100.00%	99.98%	99.98%	100.00%	100.00%	100.00%	100.00%	99.98%	100.00%
P18	99.77%	99.82%	99.79%	99.79%	99.84%	99.83%	99.82%	99.85%	99.82%	99.82%	99.87%	99.94%
Q11	99.95%	99.95%	99.94%	99.95%	99.95%	99.94%	100.00%	100.00%	100.00%	100.00%	99.99%	100.00%
R02	99.88%	99.81%	99.81%	99.81%	99.79%	99.82%	99.93%	99.88%	99.89%	99.89%	99.90%	99.90%
R03	99.40%	99.50%	99.58%	99.60%	99.46%	99.64%	99.60%	99.62%	99.75%	99.73%	99.80%	99.75%
R09	99.68%	100.00%	99.89%	99.46%	99.65%	99.64%	99.68%	100.00%	100.00%	99.67%	99.74%	99.91%
Activity weighted average	**99.71%**	**99.74%**	**99.73%**	**99.75%**	**99.75%**	**99.77%**	**99.90%**	**99.91%**	**99.90%**	**99.91%**	**99.91%**	**99.92%**

	Non-elective survival rate											
	30 days						inhospital					
HRG	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**
A07	98.66%	98.53%	98.65%	99.02%	99.01%	99.08%	99.40%	99.35%	99.63%	99.30%	99.29%	99.66%
B02	98.04%	98.79%	99.16%	98.68%	98.10%	98.49%	98.53%	99.51%	99.37%	99.83%	98.89%	99.50%
B03	98.88%	98.21%	98.15%	97.10%	100.00%	100.00%	100.00%	100.00%	98.15%	97.10%	100.00%	100.00%
C14	99.46%	99.26%	99.57%	99.47%	99.50%	99.42%	99.71%	99.54%	99.71%	99.66%	99.75%	99.53%
C22	98.86%	98.82%	98.83%	98.68%	98.89%	99.19%	99.63%	99.51%	99.41%	99.35%	99.51%	99.57%
C24	97.59%	97.41%	97.81%	97.41%	97.92%	97.97%	98.35%	98.06%	98.42%	97.98%	98.42%	98.50%
E04	93.52%	92.79%	94.41%	93.74%	95.51%	96.50%	94.11%	93.64%	94.96%	94.22%	95.90%	96.62%
E12	84.34%	84.18%	85.16%	85.43%	87.74%	88.76%	85.69%	85.71%	86.55%	86.76%	88.77%	89.64%
E15	96.52%	96.36%	97.21%	97.48%	97.86%	98.09%	97.09%	96.96%	97.69%	97.95%	98.29%	98.47%
E35	97.44%	97.29%	97.56%	97.43%	97.97%	98.14%	98.62%	98.36%	98.61%	98.42%	98.77%	98.87%
F73	95.65%	95.00%	95.82%	95.44%	95.55%	95.69%	96.62%	96.32%	96.76%	96.05%	96.46%	96.47%
F74	99.77%	99.76%	99.93%	99.86%	99.76%	99.77%	99.93%	99.86%	99.93%	99.93%	99.90%	99.94%
G01	84.85%	90.82%	91.91%	90.35%	91.46%	92.52%	85.15%	91.07%	92.56%	90.35%	91.46%	92.52%
G13	92.21%	90.20%	92.07%	92.23%	93.52%	92.61%	93.42%	91.26%	92.75%	93.07%	93.99%	92.72%
G14	99.31%	99.69%	99.59%	99.47%	99.68%	99.78%	99.47%	99.80%	99.73%	99.51%	99.79%	99.81%
H02	80.81%	78.26%	76.15%	78.98%	78.28%	82.52%	83.70%	80.55%	80.38%	82.16%	80.47%	85.56%
H04	96.25%	94.94%	96.65%	94.05%	96.14%	95.54%	98.13%	96.84%	96.65%	94.93%	96.62%	96.75%
H23	97.31%	97.06%	97.49%	97.44%	97.71%	97.68%	98.51%	98.26%	98.40%	98.51%	98.61%	98.63%
H24	99.73%	99.72%	99.63%	99.62%	99.77%	99.82%	99.87%	99.83%	99.81%	99.82%	99.88%	99.90%
H26	99.33%	99.39%	99.33%	99.38%	99.57%	99.54%	99.50%	99.52%	99.47%	99.53%	99.67%	99.69%
J01	100.00%	100.00%	96.55%	100.00%	100.00%	100.00%	100.00%	100.00%	96.55%	100.00%	100.00%	100.00%
L32	97.76%	98.22%	97.63%	98.04%	98.35%	98.43%	98.51%	99.36%	98.49%	98.81%	98.75%	99.02%
M07	99.48%	99.43%	99.44%	99.65%	99.63%	99.59%	99.62%	99.56%	99.63%	99.75%	99.71%	99.66%
M09	99.99%	99.99%	99.99%	99.99%	100.00%	99.99%	100.00%	100.00%	100.00%	99.99%	100.00%	100.00%
P18	99.54%	99.52%	100.00%	100.00%	100.00%	97.93%	99.54%	99.52%	100.00%	100.00%	100.00%	97.93%
Q11	98.09%	100.00%	99.38%	100.00%	99.04%	100.00%	99.36%	100.00%	100.00%	100.00%	99.04%	100.00%
R02	99.11%	98.56%	99.24%	98.66%	98.80%	99.32%	99.37%	99.16%	99.48%	98.91%	99.10%	99.54%
R03	93.14%	93.52%	91.88%	91.89%	95.02%	94.80%	95.86%	96.06%	95.58%	95.33%	95.83%	96.05%
R09	99.42%	98.84%	99.39%	100.00%	100.00%	98.52%	99.42%	99.42%	100.00%	100.00%	100.00%	99.01%
Activity weighted average	**94.52%**	**94.85%**	**95.37%**	**95.54%**	**96.12%**	**96.52%**	**95.26%**	**95.60%**	**96.06%**	**96.19%**	**96.65%**	**96.99%**

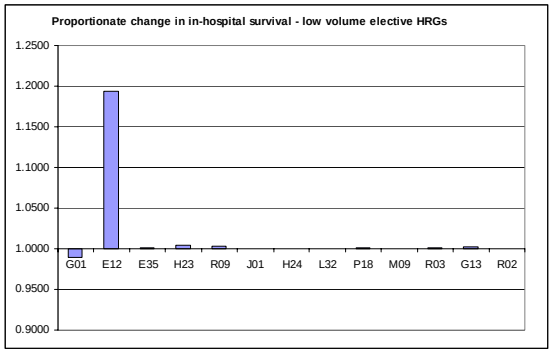
Figure 6.2 shows the change in survival rates for elective CIPS between 1998/99 and 2003/04 for these alternative measures for low and high volume HRGs. Similar information is provided for non-elective CIPS in Figure 6.3.

In general, survival rates have improved for both elective and non-elective patients. The most dramatic improvement in survival has been for non-elective acute myocardial infarction (E12), where the probability of 30-day survival increased from 85.69% in 1998/99 to 89.64% in 2003/04. As AMI is also a high volume HRG, this improvement would be expected to exert a high degree of leverage on the value of an index that included survival.

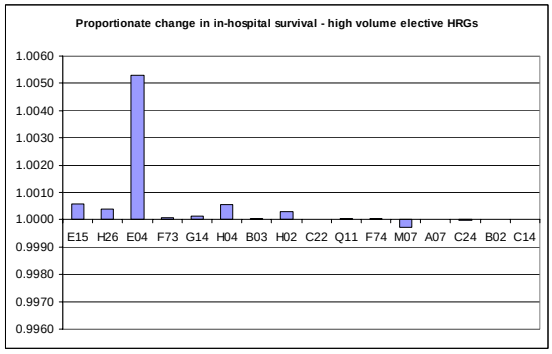
In-hospital and 30-day survival rates map each other fairly closely, as comparison of figures (a) and (c) and of figures (b) and (d) show. Consequently, it would not be expected that the index would be particularly sensitive to which measure is adopted.

Figure 6.2 Change in elective survival rates, 1998/99 – 2003/04

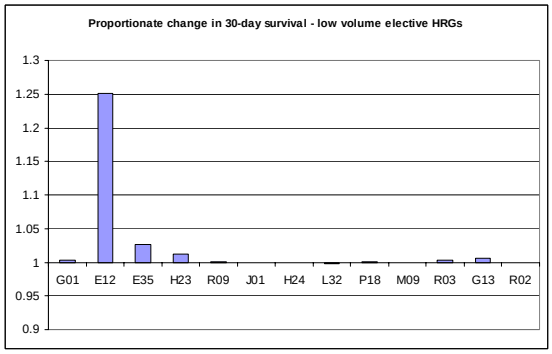
(a) In-hospital, low vol elec HRGs



(b) In-hospital, high vol elec HRGs



(c) 30 day, low vol elec HRGs



(d) 30 day, high vol elec HRGs

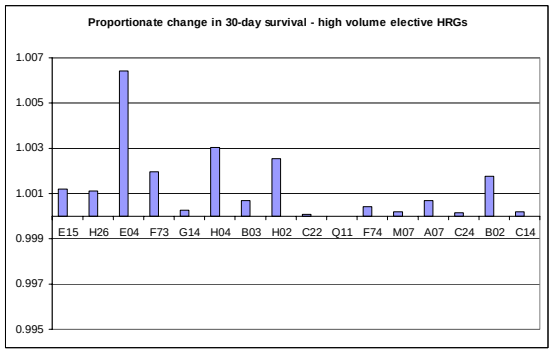
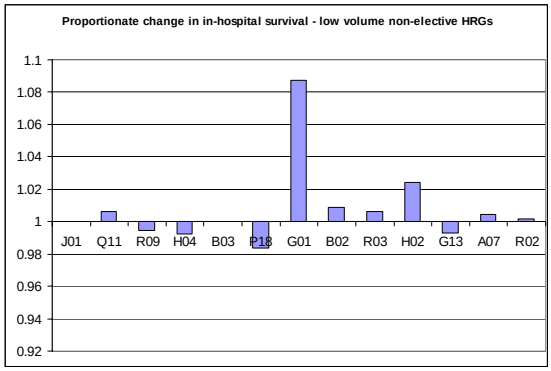
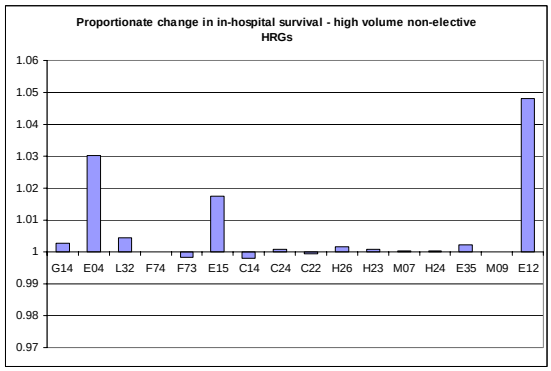


Figure 6.3 Change in non-elective survival rates, 1998/99 – 2003/04

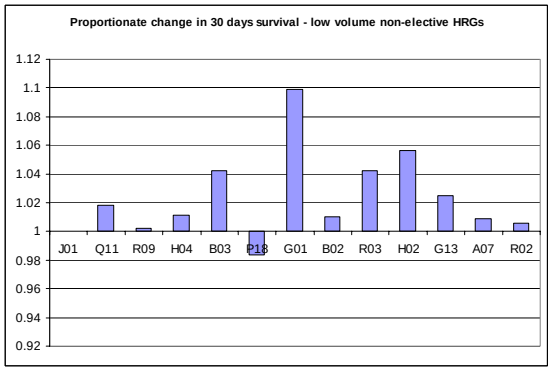
(a) In-hospital, low vol non-elec HRGs



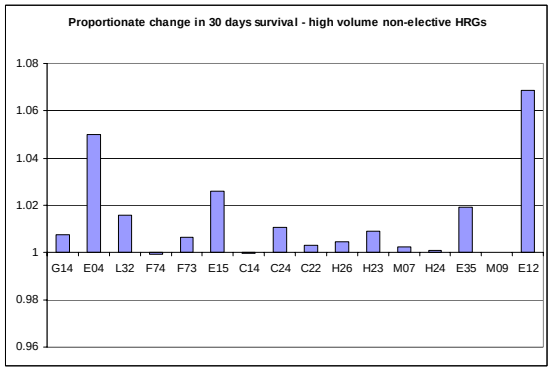
(b) In-hospital, high vol non-elec HRGs



(c) 30 day, low vol non-elec HRGs



(d) 30 day, high vol non-elec HRGs



As discussed in Appendix B, there are a variety of ways of summarising how long people have to wait for admission to hospital. In the specimen index we compare two summary measures: the mean waiting time and waiting time at the 80th percentile of the distribution. Raw data for the waiting time at the mean and 80% percentile are provided in Table 6.4. On average, the mean wait fell from 163 days to 134 days over the period, while the wait at the 80% percentile fell from 262 days to 213 days.

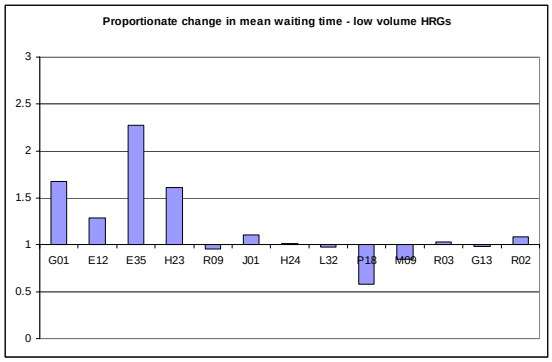
Table 6.4 Mean and 80% percentile waiting time in days, by year

	Waiting time											
	mean wait						80% percentile wait					
HRG	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**
A07	73	66	68	68	73	77	100	92	96	101	106	112
B02	228	204	193	184	180	153	357	323	310	301	295	248
B03	236	209	197	185	165	161	371	332	329	305	293	261
C14	118	97	82	75	81	90	185	143	117	106	118	142
C22	205	184	177	167	183	173	356	316	309	288	319	294
C24	144	126	109	151	128	115	245	204	167	262	211	194
E04	195	199	215	189	154	106	350	350	373	342	256	175
E12	68	40	66	60	95	88	68	40	66	60	102	91
E15	75	72	83	84	89	92	112	104	128	126	140	152
E35	37	37	40	51	77	84	37	37	40	60	83	94
F73	178	159	161	163	161	144	316	273	271	280	280	245
F74	174	151	146	147	148	139	303	250	235	237	246	232
G01	14	13	11	13	19	23	14	24	17	13	19	23
G13	147	147	151	161	161	145	252	241	265	286	290	257
G14	162	157	163	169	169	154	277	262	279	299	302	263
H02	236	238	250	253	247	225	374	383	402	404	378	340
H04	285	285	294	294	282	252	437	444	458	445	406	354
H23	58	59	56	48	67	93	65	59	56	48	71	102
H24	76	76	64	66	68	77	96	91	72	83	70	79
H26	28	26	27	28	30	35	28	26	27	28	33	43
J01	110	109	101	119	125	122	168	170	152	209	268	254
L32	72	65	68	69	79	70	86	86	91	90	122	95
M07	109	106	99	97	100	96	173	164	150	147	155	153
M09	13	12	11	13	11	11	13	12	11	13	11	11
P18	31	19	21	40	22	18	31	19	26	40	23	27
Q11	251	225	214	211	216	196	408	371	354	352	348	304
R02	110	107	120	122	127	119	173	162	173	196	228	219
R03	165	162	180	184	179	171	304	286	318	331	330	301
R09	136	130	140	143	140	130	227	216	209	234	258	247
Activity weighted average	**163**	**148**	**144**	**145**	**144**	**134**	**262**	**235**	**229**	**235**	**231**	**213**

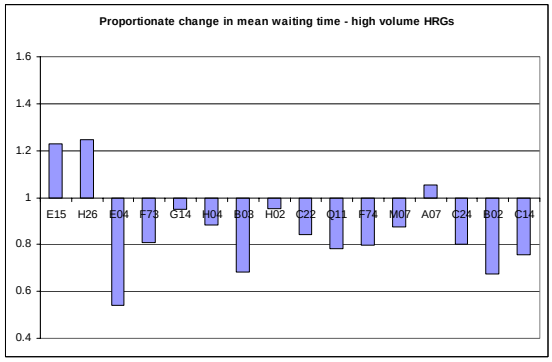
Figure 6.4 shows the change between 1998/99 and 2003/04 in the mean waiting time (Figure 6.4 (a) and (b)) and waiting time at the 80th percentile of the distribution (Figure 6.4 (c) and (d)). Although waiting times increased between 1998/99 and 2003/04 for some HRGs, these tend to be low volume activities. For the majority of high volume HRGs, waiting times fell. It would be expected that the net effect, therefore, of including waiting times in the specimen index would be an increase in the index over the period. Figures (a) and (c) and figures (b) and (d) are very similar, suggesting that the choice between mean and 80th percentile as a summary of waiting time is unlikely to have a dramatic effect on the index.

Figure 6.4 Change in waiting time, 1998/99 – 2003/04

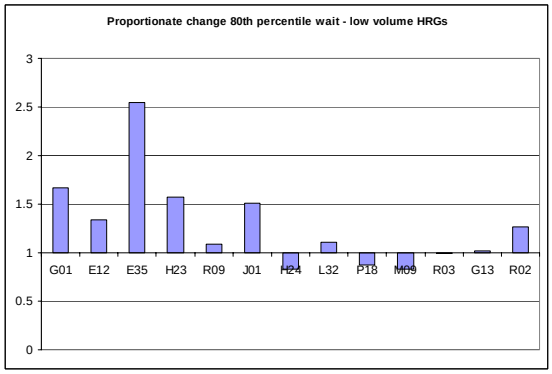
(a) Mean wait, low volume HRGs



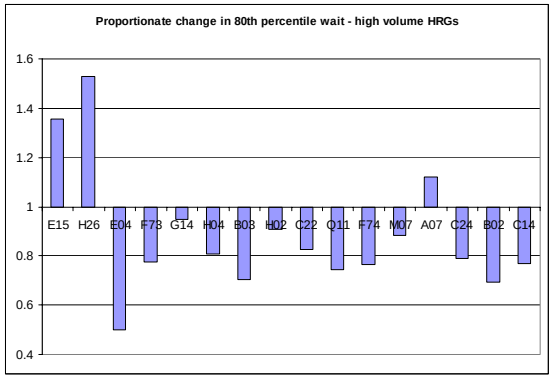
(b) Mean wait, high volume HRGs



(c) 80th perc wait, low volume HRGs



(d) 80th perc wait, high volume HRGs



As noted in section 2.7, the use of cost weights presumes an efficient allocation of NHS resources. If efficient allocation cannot be assumed, an alternative basis for establishing the relative value of activity would be according to the health outcomes each produces. The before, h_j^0, and after, h_j^*, measures of the health effect for each of the HRGs included in the specimen index are provided in Table 6.1 while unit costs, calculated on the basis of CIPS, are presented in Table 6.5. Where relative costs are not proportionate to relative health outcomes, the assumption of efficient allocation is questionable.

Table 6.5 Unit costs based on CIPS, by year

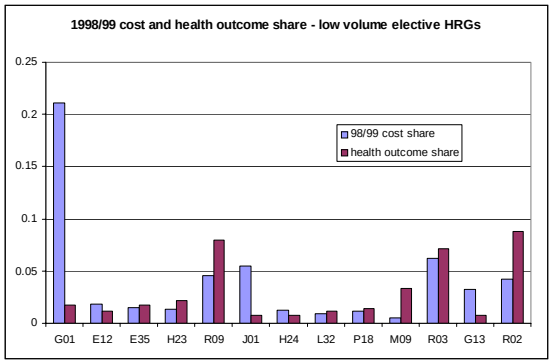
	Elective unit cost						Non-elective unit cost					
HRG	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**	**1998/99**	**1999/00**	**2000/01**	**2001/02**	**2002/03**	**2003/04**
A07	£384	£384	£395	£422	£466	£466	£1,705	£1,746	£747	£993	£1,050	£966
B02	£628	£628	£624	£673	£682	£682	£1,094	£1,096	£1,116	£1,087	£994	£1,000
B03	£661	£661	£637	£737	£734	£736	£1,299	£1,285	£1,294	£1,580	£1,112	£1,138
C14	£465	£465	£491	£550	£574	£574	£778	£785	£850	£898	£943	£930
C22	£717	£718	£741	£849	£914	£913	£979	£988	£1,171	£1,206	£1,303	£1,307
C24	£639	£641	£679	£769	£807	£806	£948	£984	£1,146	£1,256	£1,251	£1,233
E04	£5,024	£5,121	£5,654	£6,480	£6,507	£6,388	£5,264	£5,426	£5,794	£6,401	£6,973	£6,914
E12	£995	£1,224	£1,588	£1,580	£1,745	£1,687	£1,182	£1,352	£1,484	£1,688	£1,546	£1,625
E15	£2,385	£2,388	£2,421	£2,457	£2,815	£2,820	£2,587	£2,622	£2,824	£3,040	£3,241	£3,227
E35	£812	£838	£803	£1,086	£992	£942	£855	£912	£915	£970	£893	£868
F73	£898	£901	£991	£1,093	£1,148	£1,149	£1,496	£1,521	£1,795	£1,925	£1,919	£1,922
F74	£667	£667	£731	£810	£873	£872	£1,042	£1,047	£1,126	£1,327	£1,368	£1,360
G01	£11,595	£11,839	£14,149	£18,505	£18,961	£19,020	£14,569	£14,612	£18,696	£20,229	£23,179	£23,034
G13	£1,756	£1,780	£1,857	£2,055	£2,117	£2,100	£3,073	£3,202	£3,310	£3,894	£3,802	£3,750
G14	£1,304	£1,306	£1,358	£1,500	£1,564	£1,560	£1,947	£1,961	£2,081	£2,385	£2,465	£2,439
H02	£3,965	£3,993	£4,284	£4,442	£4,763	£4,758	£4,039	£4,172	£4,741	£5,191	£5,658	£5,669
H04	£4,454	£4,471	£4,661	£4,859	£5,294	£5,278	£4,655	£4,772	£4,408	£4,084	£5,477	£5,394
H23	£713	£709	£734	£861	£851	£853	£910	£966	£981	£961	£947	£936
H24	£663	£660	£625	£614	£639	£635	£654	£667	£643	£618	£614	£609
H26	£997	£1,003	£1,051	£978	£907	£902	£1,299	£1,352	£1,448	£1,521	£1,488	£1,473
J01	£3,027	£3,032	£3,346	£3,735	£3,965	£3,971	£2,947	£3,057	£3,480	£3,218	£3,340	£3,340
L32	£481	£491	£526	£628	£653	£646	£1,136	£1,217	£1,303	£1,465	£1,303	£1,285
M07	£1,839	£1,845	£1,912	£2,109	£2,298	£2,294	£1,858	£1,865	£1,818	£2,064	£2,234	£2,228
M09	£282	£282	£305	£349	£405	£405	£325	£325	£351	£400	£399	£399
P18	£641	£644	£732	£784	£783	£783	£778	£774	£589	£2,664	£1,562	£1,451
Q11	£676	£677	£727	£837	£894	£894	£1,373	£1,397	£1,319	£1,487	£1,234	£1,237
R02	£2,311	£2,332	£2,569	£2,706	£2,899	£2,896	£2,854	£2,909	£3,269	£3,649	£3,774	£3,767
R03	£3,407	£3,403	£3,727	£3,893	£4,184	£4,189	£4,813	£4,995	£5,200	£5,764	£5,961	£5,954
R09	£2,497	£2,498	£3,061	£3,178	£3,641	£3,649	£3,007	£3,075	£3,122	£3,816	£4,476	£4,406
Activity weighted average	**£1,069**	**£1,078**	**£1,150**	**£1,261**	**£1,355**	**£1,379**	**£1,077**	**£1,103**	**£1,159**	**£1,251**	**£1,274**	**£1,290**

Figure 6.5 and Figure 6.6 show the relative weight given to each HRG if relative values are based on costs or health outcomes, for elective and non-elective HRGs respectively. The figures are sub-divided to show low and high volume HRGs separately and with cost weights calculated on 1998/99 Reference Costs and 2003/04 Reference Costs. The health outcome weights are time invariant.

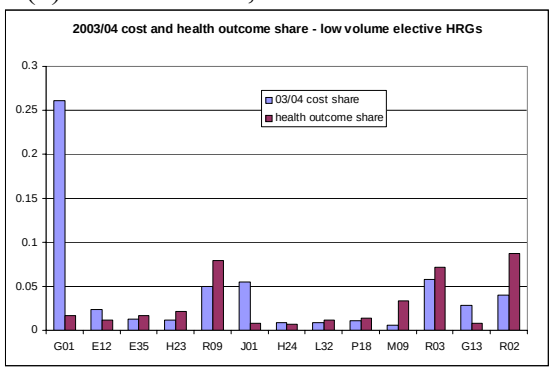
Many of the high volume HRGs appear relatively more "valuable" if value is based on health outcome rather than cost (e.g. C24, C22, H26, M09), with the stark exception of G01 (liver transplantation). All else equal, if a greater proportion of these activities were undertaken in 2003/04 compared to 1998/99, an index in which activity is valued according to health outcome would suggest greater output growth than an index where relative values are based on costs.

Figure 6.5 Cost and health outcome weights, elective HRGs

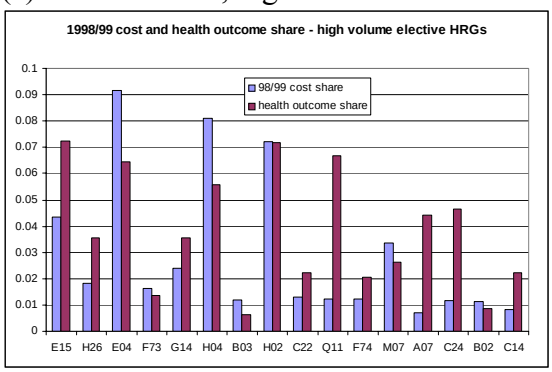
(a) 1998/99 costs, low volume HRGs



(b) 2003/04 costs, low volume HRGs



(c) 1998/99 costs, high volume HRGs



(d) 2003/04 costs, high volume HRGs

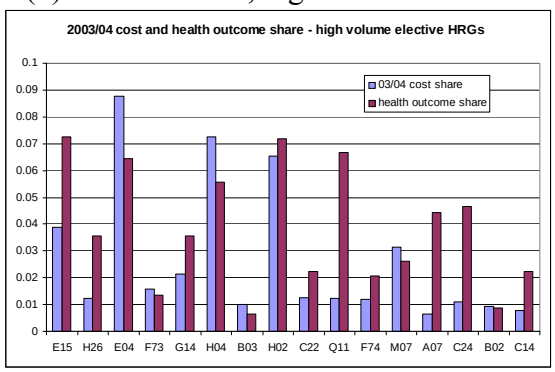
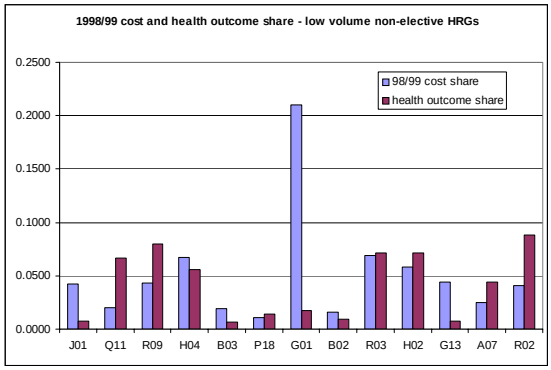
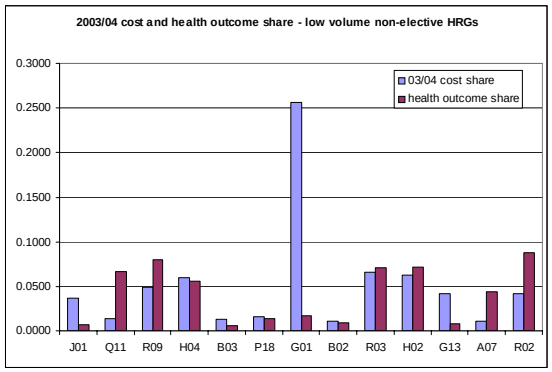


Figure 6.6 Cost and health outcome weights, non-elective HRGs

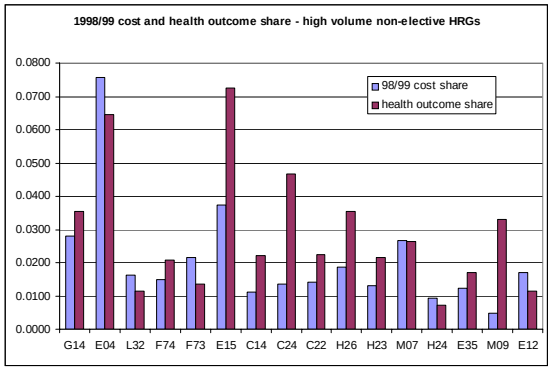
(a) 1998/99 costs, low volume HRGs



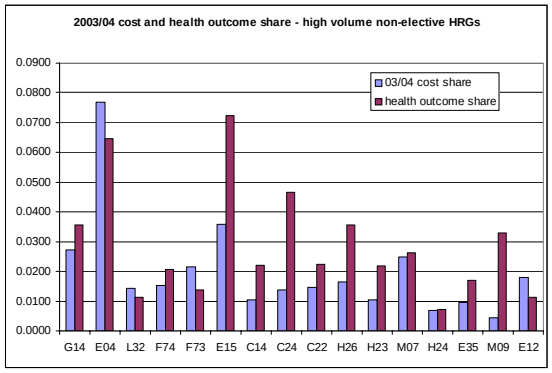
(b) 2003/04 costs, low volume HRGs



(c) 1998/99 costs, high volume HRGs



(d) 2003/04 costs, high volume HRGs



In the following sections we compare estimates of output growth under various specifications of the specimen index with CIPS measuring volume. All estimates are based on a Laspeyres index.

6.3 Cost weighted output indices

Column (i) of Table 6.6 contains estimates of output change for a cost weighted output index (CWOI) corresponding to equation (19) of the form:

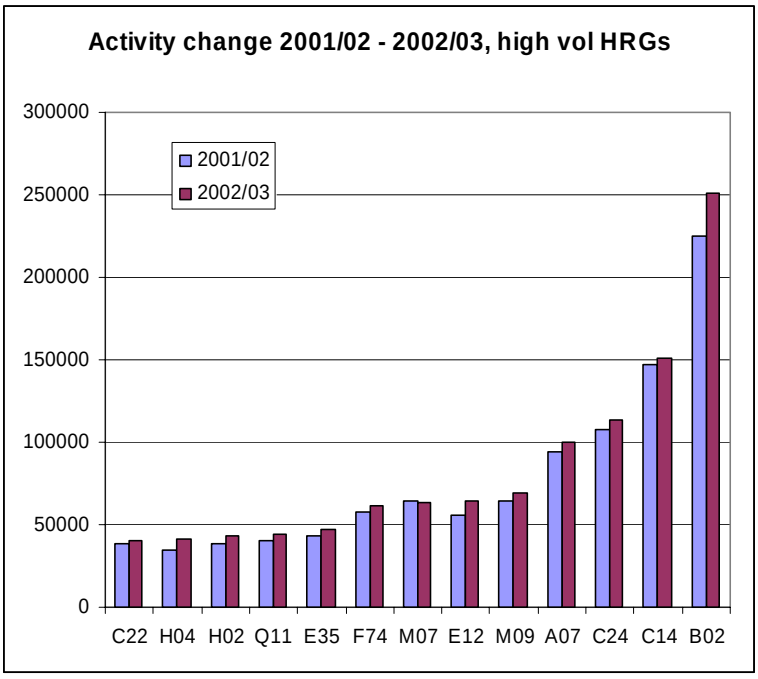
$$I_{ct}^{x} = \frac{\sum_j x_{jt+1} c_{jt}}{\sum_j x_{jt} c_{jt}}$$

Table 6.6 Cost weighted output index, with adjustments for survival and health effects

	CWOI	CWOI survival adjustn		CWOI health adjustment		
		In-hospital	30-day	30-day no threshold	30-day threshold<0.90 k=0.8	30-day
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
1998/99 - 1999/00	-1.19%	-1.18%	-1.17%	-1.33%	-1.20%	-1.19%
1999/00 - 2000/01	2.79%	2.88%	2.90%	2.66%	2.98%	3.16%
2000/01 - 2001/02	2.18%	2.20%	2.23%	1.78%	2.29%	2.57%
2001/02 - 2002/03	9.14%	9.35%	9.39%	4.76%	9.55%	12.37%
2002/03 - 2003/04	6.30%	6.42%	6.46%	6.37%	6.67%	7.49%
Average	**3.84%**	**3.93%**	**3.96%**	**2.85%**	**4.06%**	**4.88%**

This unadjusted CWOI suggests an average annual growth in output of 3.84%. There is annual variation in the estimated amount of growth. In particular, there is a large increase in the index of 9.14% between 2001/02 and 2002/03. This is driven by an increase in activity rather than a change in the costs. The average (unweighted for volume or cost) increase in activity between 2001/02 and 2002/03 was 12%. Figure 6.7 below shows the number of elective and non-elective CIPS in each year for high volume HRGs.

Figure 6.7 Activity change 2001/02-2002/03, high volume HRGs



Columns (ii) and (iii) in Table 6.6 present results for a survival adjusted cost weighted output index, of the form presented in equation (40):

$$I_{ct}^{xa} = \frac{\sum_j x_{jt+1}(a_{jt+1}/a_{jt})c_{jt}}{\sum_j x_{jt}c_{jt}}$$

Compared to the unadjusted cost weighted output index, there is slight increase in the estimated output growth when survival is included in index, reflecting the general improvement in survival over the period. The increase is slight, however, because survival rates are high for these HRGs (around 97%).

Inclusion of the survival effect increases estimated annual output growth by 0.09% if in-hospital survival is considered and 0.12% based on 30-day survival rates. Given the high correlation between these measures, their equivalent influence on the index is unsurprising. Subsequent estimations employ 30-day survival as a measure of a_j.

The figures in columns (iv) and (v) of Table 6.6 show the estimates including adjustment for before and after health status, corresponding to equation (48):

$$I_{ct}^{xq} = \frac{\sum_j x_{jt+1}(a_{jt+1}h_j^* - h_j^0)/(a_{jt}h_j^* - h_j^0)c_{jt}}{\sum_j x_{jt}c_{jt}}$$

As discussed in section 4.8.2, in estimating this equation it is necessary to introduce an arbitrary threshold for HRGs with poor survival rates. Failure to make this adjustment makes the index disproportionately sensitive to changes in a_j for activities with small or negative $(a_{jt}h_j^* - h_j^0)$ or $(a_{jt+1}h_j^* - h_j^0)$. If survival rates are below a particular threshold, only the change in survival is taken into account. To illustrate we estimate the equation without a threshold and with a threshold at the 90% survival rate.

As can be seen, omitting the threshold has a dramatic effect on the estimates, particularly in 2001/02-2002/03, where output growth was estimated as greater than 9% but now appears much lower (4.76%). The divergence stems predominantly (but not exclusively) from non-elective E12, which has a poor survival rate (of 85.69% in 1998/99). The influence of this HRG is felt particularly in the change between 2001/02-2002/03, when activity increased from 55,455 to 63,691. A formulation of the form $(a_{jt+1}h_j^* - h_j^0)$, takes a negative value for E12, with the adjustment being the ratio of two negative numbers and showing the increase in $(ah^* - h^0)$ as a reduction. This pulls the index down dramatically particularly in years where there was a growth in this activity (2001/02-2002/03) and up in years where this activity declined (e.g. 1999/00-2000/01).

Hence, for HRGs with a survival rate below 90%, the before-and-after health adjustment is not taken into account. This threshold applies in most years to E12 (AMI) and H02 (non-elective primary hip replacement) and in occasional years to G01 (liver transplantation).

The set of estimates in column (v) in Table 6.6 show estimates when the threshold is included. Inclusion of the health effects leads to an average annual increase in the estimates of output growth of 0.1% compared to the CWOI survival adjusted index.

The final set of figures (column (vi)) in table 6.6 assume that health effects are constant across treatments, ie where k=0.8. As can be seen this makes a dramatic

difference to the estimates of output growth, changing from an average of 4.06% when *k* varies by HRG to 4.88% when *k* is held constant. This sensitivity reflects both the difference in average values of *k* for the HRGs included in the specimen index ($\bar{k}_j$ =0.825) and, more particularly, the substantial variation in k_j (standard deviation=0.14). Of course, it is not possible to speculate about the direction of estimated output change from relaxing the assumption of a constant value of *k* when applied across the full range of NHS activities. However, this analysis does suggest the impact might be of substantial magnitude.

Table 6.7 presents estimates for a cost weighted output index where waiting times and life expectancy are taken into account with waiting time is discounted to date placed on the list, as described in section 4.10.2.

$$I_{ct}^{xaw} = \frac{\sum_j x_{jt+1} c_{jt} \left\{ \frac{h_j^o \left[\frac{(1-e^{-r_w w_{jt+1}})}{r_w} - \frac{(1-e^{-r_L w_{jt+1}})}{r_L} \right] + \left(a_{jt+1} h_j^* - h_j^o \right) \frac{e^{-r_L w_{jt+1}} \left(1-e^{-r_L L_{jt+1}} \right)}{r_L}}{h_j^o \left[\frac{(1-e^{-r_w w_{jt}})}{r_w} - \frac{(1-e^{-r_L w_{jt}})}{r_L} \right] + \left(a_{jt} h_j^* - h_j^o \right) \frac{e^{-r_L w_{jt}} \left(1-e^{-r_L L_{jt}} \right)}{r_L}} \right\}}{\sum_j x_{jt} c_{jt}} \quad (105)$$

Table 6.7 reports the sensitivity of results to:

- Mean and 80th percentile waiting times
- Discounting life expectancy at 1.5% or 5%
- Discounting waiting time at 1.5%, 5% or 10%

The choice between mean wait (top half of Table 6.7) and the wait experienced at the 80% percentile (bottom half of Table 6.7) has little difference on the estimates, unsurprisingly given their close correlation. Use of the 80th percentile generates slightly higher estimates of output growth, reflecting the policy concentration on reducing the waiting times for long waits during this period. In subsequent estimations the 80% percentile wait is chosen to measure waiting time.

As can be seen, the choice of a discount rate of 10% applied to waiting time has a significant effect on the estimates.

Table 6.7 Cost weighted output index, with adjustment for waiting times discounted to date placed on list

Mean waiting time	**CWOI health, waiting time and life expectancy adjustment** Waiting discounted to date placed on list					
discount rate life expectancy	1.50%			5%		
discount rate waiting time	1.50%	5%	10%	1.50%	5%	10%
1998/99 - 1999/00	-1.36%	-1.35%	-1.35%	-1.20%	-1.19%	-1.18%
1999/00 - 2000/01	2.40%	2.40%	2.39%	2.65%	2.65%	2.64%
2000/01 - 2001/02	1.96%	1.97%	1.97%	2.03%	2.03%	2.04%
2001/02 - 2002/03	9.88%	9.89%	9.90%	9.99%	9.99%	10.00%
2002/03 - 2003/04	6.58%	6.59%	6.61%	6.75%	6.77%	6.79%
Average	**3.89%**	**3.90%**	**3.90%**	**4.04%**	**4.05%**	**4.06%**
80% percentile waiting time	**CWOI health, waiting time and life expectancy adjustment** Waiting discounted to date placed on list					
discount rate life expectancy	1.50%			5%		
discount rate waiting time	1.50%	5%	10%	1.50%	5%	10%
1998/99 - 1999/00	-1.34%	-1.32%	-1.30%	-1.15%	-1.12%	-1.08%
1999/00 - 2000/01	2.40%	2.39%	2.39%	2.65%	2.65%	2.65%
2000/01 - 2001/02	1.96%	1.96%	1.97%	2.01%	2.01%	2.02%
2001/02 - 2002/03	9.92%	9.94%	9.97%	10.08%	10.11%	10.15%
2002/03 - 2003/04	6.62%	6.65%	6.70%	6.87%	6.92%	6.98%
Average	**3.91%**	**3.93%**	**3.95%**	**4.09%**	**4.11%**	**4.14%**

An alternative approach to considering waiting times is to discount to date of treatment and include a charge for waiting, as discussed in section 4.10.2.2, so that the index becomes:

$$I_{ct}^{xaw} = \frac{\sum_j x_{jt+1} c_{jt} \left(\frac{a_{jt+1} - h_j^o / h_j^*}{a_{jt} - h_j^o / h_j^*} \right) \frac{\left[\frac{\left(1 - e^{-r_L L_{jt+1}}\right)}{r_L} - \frac{\left(e^{r_w w_{jt+1}} - 1\right)}{r_w} \right]}{\left[\frac{\left(1 - e^{-r_L L_{jt}}\right)}{r_L} - \frac{\left(e^{r_w w_{jt}} - 1\right)}{r_w} \right]}}{\sum_j x_{jt} c_{jt}} \qquad (106)$$

We estimate this formulation at different discount rates, with results presented in Table 6.8. This generates higher estimates of output growth than the formulation in which waits were discounted to date placed on list.

Table 6.8 Cost weighted output index, with adjustment for waiting times discounted to date of treatment

80% percentile waiting time	CWOI health, waiting time and life expectancy adjustment Waiting discounted to date of treatment, with charge for waiting							
discount rate life expectancy	1.50%	1.50%			5%			1.50%
discount rate waiting time	1.50%	1.50%	5%	10%	1.50%	5%	10%	0%
	k=0.8							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
1998/99 - 1999/00	-1.17%	-1.18%	-1.17%	-1.16%	-0.99%	-0.98%	-0.97%	-1.39%
1999/00 - 2000/01	2.50%	2.32%	2.32%	2.31%	2.58%	2.57%	2.57%	2.40%
2000/01 - 2001/02	2.33%	2.05%	2.05%	2.06%	2.11%	2.11%	2.12%	2.00%
2001/02 - 2002/03	12.46%	9.65%	9.67%	9.69%	9.81%	9.83%	9.86%	9.27%
2002/03 - 2003/04	8.10%	7.29%	7.32%	7.35%	7.52%	7.55%	7.59%	6.54%
Average	**4.85%**	**4.03%**	**4.04%**	**4.05%**	**4.20%**	**4.22%**	**4.23%**	**3.76%**

Column (viii) has the results of setting r_w = 0 so that there is no adjustment for waiting, only for life expectancy as in section 4.8.3. Compared to the CWOI with only a health effects adjustment Table 6.6, column (v) the average growth rate is reduced by about 0.3%.

6.4 Health outcome weighted output indices

This section presents the results from calculation of health outcomes weighted output indices (HOWOI). For comparative purposes with the CWOI, we first estimate a version of HOWOI in which the impact of treatment on life expectancy is ignored. In effect, this amounts to comparing the use of cost and survival-adjusted before and after health outcomes (not QALYs) as weights:

$$\frac{\sum_j x_{jt+1}(a_{jt+1}h_j^* - h_j^0)}{\sum_j x_{jt}(a_{jt}h_j^* - h_j^0)} \qquad (107)$$

This equation is estimated both with k varying by HRG (column (ii) Table 6.9) and for a value of k=0 (column (iii) Table 6.9). As can be seen, output growth appears lower in this formulation of a HOWOI than the corresponding cost weighted output index (figures from Table 6.6 reproduced in column (i) of Table 6.9).

As can be seen, substitution of cost for health outcome weights leads to a reduction in estimated output growth for this group of HRGs. The extent to which an index is sensitive to the choice of cost and health outcome weights depends on three factors:

- Whether cost weights are disproportionate to health outcome weights;

- The volume of activity in those HRGs where the relative weights are most disproportionate;
- The change in activity over time in those HRGs where the relative weights are most disproportionate.

For a handful of HRGs, cost weights are greater than health outcome weights. This is particular evident for G01 liver transplants, which are costly (the non-elective cost was £23,000 in 2003/04) but their estimated contribution to health outcome is about average for the sample of HRGs considered here. However, because this is a low volume HRG and there is little change in the amount of activity over time, the impact of changing the valuation basis for G01 exerts little influence on the overall index.

In contrast, elective activity categorised to A07 (intermediate pain procedures) contributes 7.4% of total 2003/04 activity, and elective activity in this HRG grew by 17.5% over the period captured by the index. Its cost share in 2003/04 is only 0.6% whereas its health outcome share is 4.43%. All else equal, the growth in activity in this high volume HRG would lead to an index based on health outcome shares having a higher value than one based on cost shares.

There is little difference between the cost and health outcome weights for B02 (cataract extractions), but they are accorded slightly less weight (-0.06%) when relative values are based on health outcomes. However, despite this minimal difference, B02 exerts considerable influence on the overall index, contributing 20.7% of total volume in 2003/04. There has also been a volume increase of 81% in this activity over the period captured by the index. Thus, this HRG exerts downward influences on the index.

Table 6.9 Health outcome weighted output index

	CWOI	HOWOI		HOWOI	
		No LE	No LE k=0.8	With LE	
discount rate				1.50%	5%
life expectancy					
	(i)	(ii)	(iii)	(iv)	(v)
1998/99 - 1999/00	-1.19%	-2.96%	-1.88%	-4.75%	-4.06%
1999/00 - 2000/01	2.79%	0.16%	1.87%	-3.78%	-2.20%
2000/01 - 2001/02	2.18%	1.41%	1.04%	0.27%	0.48%
2001/02 - 2002/03	9.14%	10.17%	9.11%	8.43%	8.85%
2002/03 - 2003/04	6.30%	4.62%	5.07%	1.95%	2.73%
Average	**3.84%**	**2.68%**	**3.04%**	**0.42%**	**1.16%**

The previous adjustment makes the assumption that the health status snapshots h_j^*, h_j^o measure the discounted sum of QALYs q_j^*, q_j^o. More properly health outcome weights should incorporate the effect of treatment on life expectancy, so that they more nearly measure the discounted sum of QALYs:

$$\frac{\sum_j x_{jt+1}(a_{jt+1}h_j^* - h_j^0)(1-e^{-r_L L_{jt+1}})}{\sum_j x_{jt}(a_{jt}h_j^* - h_j^0)(1-e^{-r_L L_{jt}})} \qquad (108)$$

Estimates are presented from this index in columns (iv) and (v) of table 6.9, with life expectancy discounted at 1.5% and 5%. The impact of including life expectancy is a substantial reduction in estimated output growth. The reason for this is that life expectancy declined gradually over the period, the main reason for this probably being that increasingly older people were receiving treatment, as demonstrated in column. Table 6.10 provides evidence.

Table 6.10 Average age and life expectancy

	Age	Life expectancy
1998/99	45.71	25.83
1999/00	45.76	25.34
2000/01	46.17	24.12
2001/02	46.3	23.91
2002/03	46.93	23.59
2003/04	47.9	22.98

The health outcomes weighted output index is also estimated after incorporating

waiting times to date placed on list

$$\frac{\sum_j x_{jt+1}\left\{h_j^o\left[\frac{(1-e^{-r_w w_{jt+1}})}{r_w}-\frac{(1-e^{-r_L w_{jt+1}})}{r_L}\right]+\left(a_{jt+1}h_j^*-h^o\right)\left[\frac{e^{-r_L w_{jt+1}}\left(1-e^{-r_L L_{jt+1}}\right)}{r_L}\right]\right\}}{\sum_j x_{jt}\left\{h_j^o\left[\frac{(1-e^{-r_w w_{jt}})}{r_w}-\frac{(1-e^{-r_L w_{jt}})}{r_L}\right]+\left(a_{jt}h_j^*-h^o\right)\left[\frac{e^{-r_L w_{jt}}\left(1-e^{-r_L L_{jt}}\right)}{r_L}\right]\right\}} \quad (109)$$

and to date of treatment with a charge for waiting:

$$\frac{\sum_j x_{jt+1}\left(a_{jt+1}h_j^*-h_j^o\right)\left[\frac{\left(1-e^{-r_L L_{jt+1}}\right)}{r_L}-\frac{\left(e^{r_w w_{jt+1}}-1\right)}{r_w}\right]}{\sum_j x_{jt}\left(a_{jt}h_j^*-h_j^o\right)\left[\frac{\left(1-e^{-r_L L_{jt}}\right)}{r_L}-\frac{\left(e^{r_w w_{jt}}-1\right)}{r_w}\right]} \quad (110)$$

The sensitivity of these two variants of the HOWOI are assessed with respect to:

- Discounting life expectancy at 1.5% or 5%
- Discounting waiting time at 1.5%, 5% or 10%

Results are provided in Table 6.11. When waiting time is discounted to date placed on list, estimates of output growth are slightly higher than those when no waiting time adjustment is made, with the difference increasing at higher discount rates. Compared to discounting to date placed on list, discounting to date of treatment results in lower estimates of output growth, decreasing at higher discount rates.

Table 6.11 HOWOI, adjusted for waiting time

80% percentile waiting time	HOWOI health, waiting time and life expectancy adjustment					
	Waiting discounted to date placed on list					
discount rate life expectancy	1.50%			5%		
discount rate waiting time	1.50%	5%	10%	1.50%	5%	10%
1998/99 - 1999/00	-4.64%	-4.60%	-4.54%	-3.79%	-3.71%	-3.61%
1999/00 - 2000/01	-3.74%	-3.74%	-3.73%	-2.08%	-2.07%	-2.06%
2000/01 - 2001/02	0.20%	0.19%	0.18%	0.29%	0.28%	0.26%
2001/02 - 2002/03	8.47%	8.48%	8.50%	8.98%	9.00%	9.03%
2002/03 - 2003/04	2.00%	2.03%	2.08%	2.86%	2.91%	2.99%
Average	**0.46%**	**0.47%**	**0.50%**	**1.25%**	**1.28%**	**1.32%**
80% percentile waiting time	**HOWOI health, waiting time and life expectancy adjustment**					
	Waiting discounted to date of treatment, charge for waiting					
discount rate life expectancy	1.50%			5%		
discount rate waiting time	1.50%	5%	10%	1.50%	5%	10%
1998/99 - 1999/00	-4.52%	-4.52%	-4.51%	-3.66%	-3.65%	-3.63%
1999/00 - 2000/01	-3.81%	-3.81%	-3.82%	-2.18%	-2.18%	-2.18%
2000/01 - 2001/02	0.09%	0.09%	0.08%	0.21%	0.20%	0.19%
2001/02 - 2002/03	8.55%	8.55%	8.56%	9.06%	9.06%	9.07%
2002/03 - 2003/04	2.11%	2.11%	2.12%	3.01%	3.02%	3.04%
Average	**0.48%**	**0.48%**	**0.49%**	**1.29%**	**1.29%**	**1.30%**

6.5 Value weighted output index

Our "ideal" index takes the form specified in equation (12), in which activities are valued according to their associated health outcomes and waiting times are considered a characteristic of health care. The index takes the form:

$$I_{yt}^{xq} = \frac{\sum_j x_{jt+1}[(a_{jt+1}h_j^* - h_j^0)[(1-e^{-r_L L_{t+1}})\pi_h]/r_L - w_{jt+1}\pi_W]}{\sum_j x_{jt}[(a_{jt}h_j^* - h_j^0)(1-e^{-r_L L_t})\pi_h]/r_L - w_j\pi_W]} \qquad (111)$$

We estimate this index assuming that the monetary value of a QALY (π_h) in 2002/3 is £30,000 and applying growth rates in money GDP to calculate values for earlier years. We explore the sensitivity of results to:

- the cost of a day spent waiting (π_W) - either £3.13 or £50 in 2002/3 (adjusted by money GDP growth in earlier years)
- discounting life expectancy at 1.5% and 5%

Estimates of output growth are presented in Table 6.12. These imply lower rates of output growth than a CWOI for these HRGs, the main reason being because of the

influence of treating an increasing older population (leading to decreasing life expectancy). The effect of applying a higher value to the cost of a day spent waiting is to increase estimated output growth, but not substantially.

Table 6.12 Value weighted output index

	Value weighted output index			
cost of day spent waiting	£3.13		£50	
discount rate life expectancy	1.50%	5%	1.50%	5%
	(i)	(ii)	(iii)	(iv)
1998/99 - 1999/00	-4.71%	-4.00%	-3.39%	-1.54%
1999/00 - 2000/01	-3.86%	-2.32%	-4.22%	-2.53%
2000/01 - 2001/02	0.23%	0.41%	-0.29%	-0.42%
2001/02 - 2002/03	8.41%	8.82%	8.27%	8.70%
2002/03 - 2003/04	2.00%	2.82%	2.90%	4.59%
Average	**0.41%**	**1.15%**	**0.65%**	**1.76%**

6.6 Conclusion

In this section we have applied various formulations of an output index to a limited set of HRGs for which data were available on health status before and after treatment. The main conclusions are that:

- Estimates of output growth are sensitive to whether k is assumed constant across treatments. In view of this, it would be advisable to ascertain before and after health status for a larger sample of NHS treatments.
- There is a high correlation between indices using the two mortality measures - in-hospital and 30-day survival.
- Although relative cost and health outcome weights differ to some extent for our specimen set of HRGs, the difference does not lead to dramatic changes in the estimates produced by the specimen index. It cannot be assumed, however, that there will not be greater divergence between indices using costs and health outcome weights for other NHS activities.
- Unable to estimate QALYs directly, we have had to rely upon life tables from the general population to generate estimates of life expectancy. With an increasingly older population being treated over time, this leads to decreasing life expectancy, which in turn implies declining output growth in indices

where life expectancy is included. This is because our index formulations make the value judgement that an additional quality adjusted life year should have the same value whatever the age of the person it accrues to.

- Cost weighted indices with waiting time adjustments are sensitive to whether waiting time is discounted to the date placed on the list or to the date of treatment, and to the choice of discount rate.
- The health and waiting time outcomes index, for the HRGs considered here, is not particularly sensitive to which point in the distribution is chosen to measure waiting time (mean or 80% percentile) or to the cost applied to a day spent waiting (£3.13 or £50).

7 Effects of quality adjustments on hospital and NHS output indices: summary

In Section 4 we argued that it was important to include estimates of health effects in a quality adjusted output index. This should be done by regular collection of health outcomes data for a representative range of NHS activity. In the absence of this data, in Section 5 we used available information on outcomes for 29 HRGs and made the assumption that the average health gain observed could be applied uniformly to all hospital activity.

For the specimen quality adjusted output index discussed in Section 6, it was possible to test the sensitivity of results to the assumption of a uniform effect. For the specimen index we were able to estimate quality adjusted output using data for actual health effects and compare the result with estimates using a uniform health effect. As expected, the move from uniform to actual values does affect the result.

We recommend that wherever possible actual health effects data be used to estimate quality adjusted output indices. Over the next few years the number of HRGs for which actual data will be available should increase. This will gradually reduce the proportion of activity where it is necessary to make assumptions about health effects.

A consequence of this recommendation is that for the next few years a quality adjusted output index would have to be based on a mix of actual and assumed values.

In this section we examine the impact of departing from the assumption of uniform fixed health effect (k = 0.8) and instead use actual values where they exist and assumed values where data is absent.

- For the 29 elective procedures for which we have data, k varies by HRG as in the specimen index.
- For all other elective procedures we assume k = 0.8 as suggested by the mean of the k for the elective HRGs where there are estimates
- For non-elective HRGs we assume k = 0.4 on the grounds that non-elective patients may have worse health (q^o) if not treated so that the ratio of health if not treated to health if treated ($k = q^o/q^*$) is smaller.

Given that non-elective activity is growing more rapidly than elective, the lack of knowledge of health state and health gain for non-elective patients is a serious problem.

We compare our recommended variant (Q2) based on a health effects adjustment which varies by HRG with a variant (Q1) with the same health effect adjustment for all HRGs, elective and non-elective.

Quality variant 1 assumes $k = q^0/q^* = 0.8$ if $a - k > 0.05$ and $k = 0$ otherwise for all elective and non-elective HRGs, discounts to date of treatment with charge for wait, with discount rates on waits and health equal to 1.5% and the waiting time variable is the 80th percentile wait in each HRG.

Quality variant 2 is our recommended quality variant. This sets $k = q^0/q^* = 0.8$ for electives, k = 0.4 for non-electives, k = actual k for those HRGs included in the specimen index where this is known, provided $a - k > 0.10$ and $k = 0$ otherwise. This quality variant discounts to date of treatment with charge for wait, with discount rates on waits and health equal to 1.5% and uses 80th percentile waits.

We show the effects of these two quality adjustments variants on the HES hospital output in Table 7.1, to all hospital output including outpatients and accident and emergency in Table 7.2 and to all NHS output in Table 7.3.

Table 7.1 shows that the Q1 quality adjustment variant, with survival, health effects, life expectancy and waiting time adjustments adds just under one percentage point to the HES hospital unadjusted index average across the five growth periods. The recommended variant Q2 results in a smaller upward adjustment of just over 0.5%

Table 7.1 HES hospital cost weighted output index with hospital sector quality adjustments

	Unadjusted	Quality variant 1		Quality variant 2	
		Survival and health effect only	Survival, health effect, life expectancy and waiting	Survival and health effect only	Survival, health effect, life expectancy and waiting
1998/99-1999/00	1.87	0.09	0.49	0.63	1.04
1999/00-2000/01	0.91	1.97	1.73	1.50	1.25
2000/01-2001/02	0.95	1.01	0.87	0.82	0.65
2001/02-2002/03	4.44	7.72	7.48	6.77	6.52
2002/03-2003/04	5.81	8.04	8.15	7.21	7.31
Average	2.80	3.77	3.74	3.38	3.35

Table 7.2 shows the effect of the two variants on a broader definition of hospital activity. All variants have faster growth than for the narrower HES output indices in Table 7.1. The main reason for this is the faster growth in the activities not captured by HES, such as A&E and outpatients, which are excluded from Table 7.1. However, because fewer of the quality adjustments apply to these non-HES activities the proportionate effect of Q1 and Q2 is smaller than in Table 7.1. Thus Q1 increases average annual growth by 0.44% instead of nearly 1% and Q2 increases growth by 0.25% instead of 0.55% over the period.

Table 7.2 Hospital sector cost weighted output index with hospital sector quality adjustments

	Unadjusted	Quality variant 1		Quality variant 2	
		Survival and health effect only	Survival, health effect, life expectancy and waiting	Survival and health effect only	Survival, health effect, life expectancy and waiting
1998/99-1999/00	2.03	0.43	0.79	0.91	1.28
1999/00-2000/01	1.54	2.35	2.16	1.99	1.80
2000/01-2001/02	4.48	4.52	4.43	4.40	4.31
2001/02-2002/03	3.94	5.71	5.57	5.19	5.06
2002/03-2003/04	4.78	5.94	6.00	5.51	5.56
Average	3.35	3.79	3.79	3.60	3.60

Finally Table 7.3 shows the impact of the quality adjustments to the hospital sector output on the cost weighted output index for the NHS as a whole. We first show the CWOI without quality adjustments and then add variants of the adjustments for survival and waiting times for the hospital sector. Overall NHS output growth is higher than either of the hospital sector output growth rates because of the more rapid growth in some non-hospital activities such as prescribing and consultation rates, and because of increasing coverage of NHS activity. Quality adjustment variant Q1 increases average annual growth by 0.29% and Q2 increases it by 0.17%. Notice that for 2001/02 to 2002/03 both variants have a much larger effect (1.04% for Q1 and 0.71% for Q2) but rather small effects in the middle years and actually reduce growth from 1998/99 to 1999/00. This negative adjustment is due, as we noted in section 5.4.1, to the fall in survival for a small number of high activity high cost HRG. Notice also that from 1998/99 to 1999/00 when both Q1 and Q2 lead to downward adjustments our recommended variant Q2 has a smaller negative effect. Thus in general variant Q2 has a smaller positive or negative effect than Q1 because it uses smaller assumed health effects for emergency activities.

Table 7.3 Aggregate NHS cost weighted output index with hospital sector quality adjustments

	Unadjusted	Quality variant 1		Quality variant 2	
		Survival and health effect only	Survival, health effect, life expectancy and waiting	Survival and health effect only	Survival, health effect, life expectancy and waiting
1998/99-1999/00	2.61	1.77	1.96	2.03	2.22
1999/00-2000/01	2.11	2.57	2.46	2.36	2.26
2000/01-2001/02	3.85	3.88	3.82	3.80	3.74
2001/02-2002/03	5.07	6.20	6.11	5.87	5.78
2002/03-2003/04	4.43	5.17	5.20	4.89	4.93
Average	3.62	3.92	3.91	3.79	3.79

We next examine input changes over the period and then in section 9 combine our output indices and input indices to calculate productivity growth rates.

8 Labour input

8.1 Introduction

Current practice by DH and ONS calculates labour input by deflating payments to labour by a wage index. It is more usual to estimate labour inputs based on number of workers or hours worked so it was considered useful to devote effort in the project to this alternative method of measuring labour input. In addition the Atkinson Report recommended that labour input should be adjusted to take account of variations in types of workers employed, in particular the changing use of skilled workers; this section also addresses this recommendation. Labour is by far the most important input used in producing health services, accounting for about 75% of total hospital expenditures. These measures can then be combined with the aggregate and hospital output measures given in section 7 to calculate labour productivity growth rates and combined with measures of payment to labour can be used to calculate total factor productivity growth rates. Productivity estimates are presented in the next section.

8.2 Labour input in the NHS

This section summarises results on constructing measures of labour input.

8.2.1 Volume of labour input

The volume of labour input can be calculated using direct or indirect measures. Direct measures include number of persons engaged (including self-employed), number of full-time equivalents or total hours worked. The indirect measure is expenditure on labour deflated by a wage index, employed by ONS for NHS labour input. Productivity analysts tend to prefer direct measures since reasonable data are generally available on numbers employed whereas wage indexes are seen as less reliable. This section follows this tradition of using direct measures. However it should be noted that in sectors where the self-employed account for a large share of employment, as is the case for the NHS, there may be more grounds for using an indirect measure; this is discussed further below.

The simplest direct measure is a headcount of number of persons employed. Since many persons in the NHS work part-time, a more reliable indicator is full-time equivalent workers. Such a measure is calculated by the DH where part-time work is weighted by normal weekly hours of these people. While full-time equivalents is undoubtedly a better measure than headcounts, it is only a half-way house to the measure recommended by the OECD productivity manual (OECD, 2001) of annual actual hours worked. Normal or usual hours worked do not take account of changes in time lost due to holidays, sickness etc. Over time trends in time paid but not worked tend to dominate changes in usual weekly hours worked. Adjustments to an annual actual hours worked basis are discussed below.

8.2.2 Quality of labour input

> *"Because a worker's contribution to the production process consists of his/her "raw" labour (or physical presence) and services from his/her human capital, one hour worked by one person does not constitute the same amount of labour input as one hour worked by another person*", OECD productivity manual (p. 41).

Volume measures of labour input hide considerable diversity across types of workers. Obviously the productivity of highly skilled workers is greater than that of less skilled workers as set out in the quote above. Division by skill type is not the only quality dimension; other candidates are age or experience, gender or occupation. Nevertheless most research on measuring labour quality suggests skill is the most important dimension (Jorgenson, Ho and Stiroh, 2005).

The standard growth accounting formula for adjusting for skills divides labour hours by skill type and then weights the growth in hours of each type by their wage bill shares. This captures the fact that more highly skilled workers get paid more than the unskilled, and under competitive market conditions, the wage paid reflects the marginal productivity of workers of different types. Merely calculating growth in total hours worked is equivalent to weighting worker types by their share in employment. Hence if there is general upskilling of the workforce so that growth in hours is greater for skilled relative to unskilled workers, weighting by wage bill shares leads to higher aggregate labour input growth.

Formally, quality adjusted labour input, with *s* types of skilled labour can be calculated using a Törnqvist index (see section 3) by:

$$\ln L_t - \ln L_{t-1} = \sum_l \varpi^s (\ln L_{st} - \ln L_{st-1}) \qquad (112)$$

where *Ls* is number of hours worked for worker of skill type *s*, and

$$\varpi^s = 0.5\left(\frac{w_{s,t-1}L_{s,t-1}}{w_{t-1}L_{t-1}} + \frac{w_{st}L_{st}}{w_t L_t}\right)$$

is the wage bill share of type *s* workers in the total wage bill for all workers, averaged across periods *t* and *t*-1. The difference between the equation above and the growth in total hours worked gives the impact of skills on aggregate labour input growth.

8.2.3 Data sources and volume trends

This report reviewed available data sources relating to health sector labour input to assess their usefulness in constructing direct measures of labour input for the NHS. Two sources seemed particularly useful and formed the basis of the calculations

presented in this section. These were:

- **NHS Workforce Census** – An annual census conducted by the Department of Health. This source provides data on numbers employed in the NHS, both in headcount and full-time equivalent terms, by occupation and organisation.

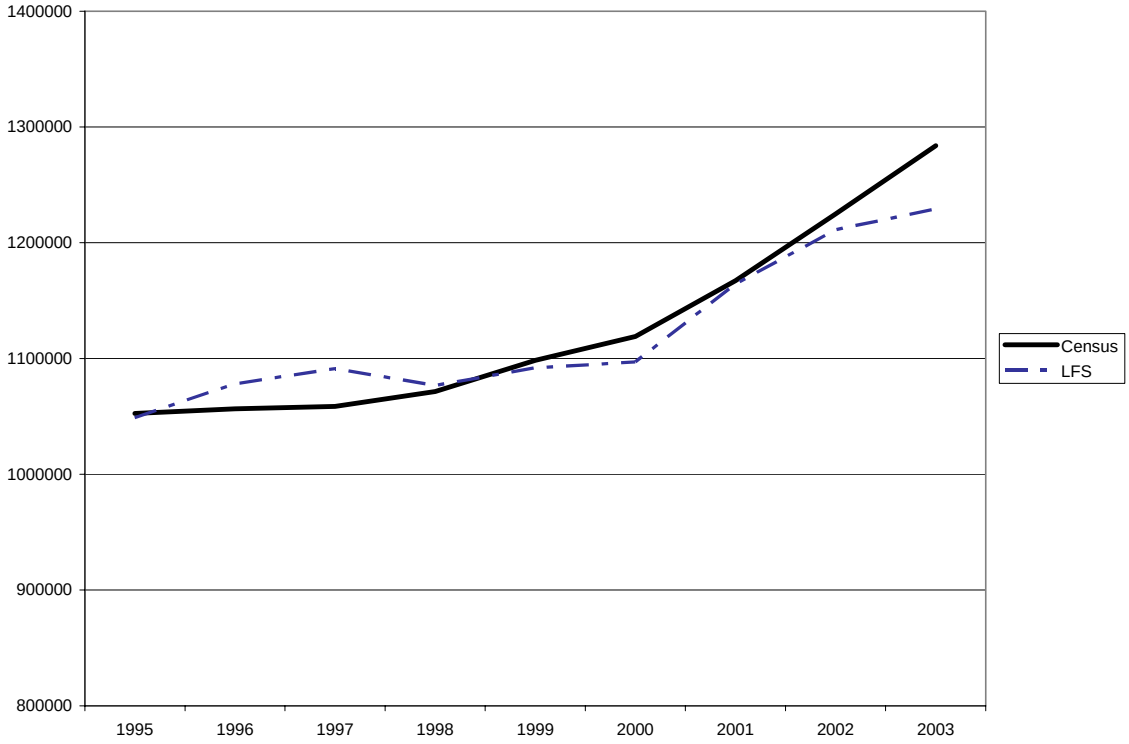
- **Labour Force Survey** – This is a quarterly sample survey conducted by ONS. It contains data on numbers employed and annual hours worked by industry (SIC92) and occupation (SOC), distinguishing private and public sectors and whether the person is employed by an NHS Trust. It also contains data on wages, qualifications, region and nationality of employees and questions relating to on the job training.[10]

These sources were supplemented by information from the NHS staff earnings survey.

The NHS Workforce Census has an advantage over the Labour Force Survey in that being a census it captures all people whose employer is the NHS. The labour force survey headcounts are derived indirectly through both the industry in which the individual states they are working (SIC 85.1 – human health activities) and if they state that they work for an NHS Trust. Against this the LFS includes agency workers whereas the NHS Census excludes these. The two sources were compared for consistency and were found to follow similar but not identical trends through time (shown in Figure 8.1).

[10] Note since the LFS is an individual survey, where the person interviewed is frequently reporting for their spouses or partner, reporting errors can lead to situations where individuals state that they work for industries such as personal services but that they are employed by the NHS, e.g. contract cleaners

Figure 8.1 Numbers employed in the NHS 1995-2003



The decision was made therefore to employ the NHS Census data for the headcount but to include an adjustment using the ratio of agency to other staff from the LFS to adjust the Census data. The following table shows growth rates in persons engaged by broad occupational group. These show significant growth in total numbers employed, in particular since 2000. The growth rates are fairly uniform across broad categories. Within categories there is more variation. For example the number of managers, included in the infrastructure support group, increased by about 6.5% per annum from 1995 to 2003, with a very large increase of over 11% per annum since 2000. However it should be noted that managers represent a very small proportion of the NHS workforce, reaching only 2.7% by 2003 despite the high growth. In contrast nursing assistants and auxiliaries show growth of less than 2% p.a. since 1995 and 2.7% since 2000.

Table 8.1 Trends in the NHS workforce

	1995-03	2000-03
Total	2.49	4.58
Professionally qualified clinical staff	2.77	4.53
All doctors	3.30	4.03
All qualified nurses (including practice nurses)	2.48	4.66
Total qualified scientific, therapeutic & technical staff	3.60	4.73
NHS infrastructure support	1.26	4.66
Support to Clinical Staff	3.20	5.35
Other	0.55	1.98

Rather than use full-time equivalents we employ LFS data on weekly hours to convert these headcounts to total annual hours. Weekly actual hours show a slight decrease over time from 28.9 in 1995 to 28.6 in 2003 with rises in between.

8.2.4 Quality adjustments based on qualifications

The LFS is used in this project to incorporate quality adjustments. Thus the NHS Census data are used as control totals with proportions of workers in each skill group and their wage rates from the LFS used in the skills adjustment. We then refine these estimates in a number of ways to take account of on the job training, regional variations in wage rates and country of birth of workers. In addition we include an adjustment for doctors using data from the NHS Census and the Earnings survey since the qualification division in the LFS is not fine enough to ensure we are picking up all skill variations.

Table 8.2 shows the proportion of workers by qualification group for selected years. Employment growth has been relatively strong among workers with higher degrees and primary degrees as well as those with A-levels and equivalents. The share of those with nursing qualifications or other NVQ4 has declined reflecting the growth in degrees among nurses and health care professionals. There has also been a marked

decline in the share of the workforce with no skills. While the figures confirm the well known high growth in university degrees among doctors, nurses and other health professionals, the changes in the lower end of the skill distribution suggest upskilling also among other NHS employees, illustrated by the chart for health care assistants below. Interestingly, managers have also experienced pronounced changes in their skill distribution with the percent of managers having a higher degree (mostly masters degrees) rising from about 3% in 1995 to about 16% in 2003 and the percent with degrees rising from 22% to 28% over the same time period (Figure 8.2).

Table 8.2 Skill proportions of the Workforce: NHS selected years

	1995	2000	2003
Higher degree (Masters, PhDs)	5.4	6.9	8.3
Degree	13.2	17.4	18.2
Nursing qualification/other NVQ4 (higher education below degree)	36.4	31.3	29.2
NVQ3 (A-levels or equivalent)	6.1	8.6	9.8
NVQ1 and NVQ2 (GCSEs or equivalent, vocational qualifications)	22.8	23.8	23.2
Other	6.8	5.9	5.9
None	10.0	6.1	5.5

Figure 8.2 Healthcare assistants: proportions employed by highest qualification held

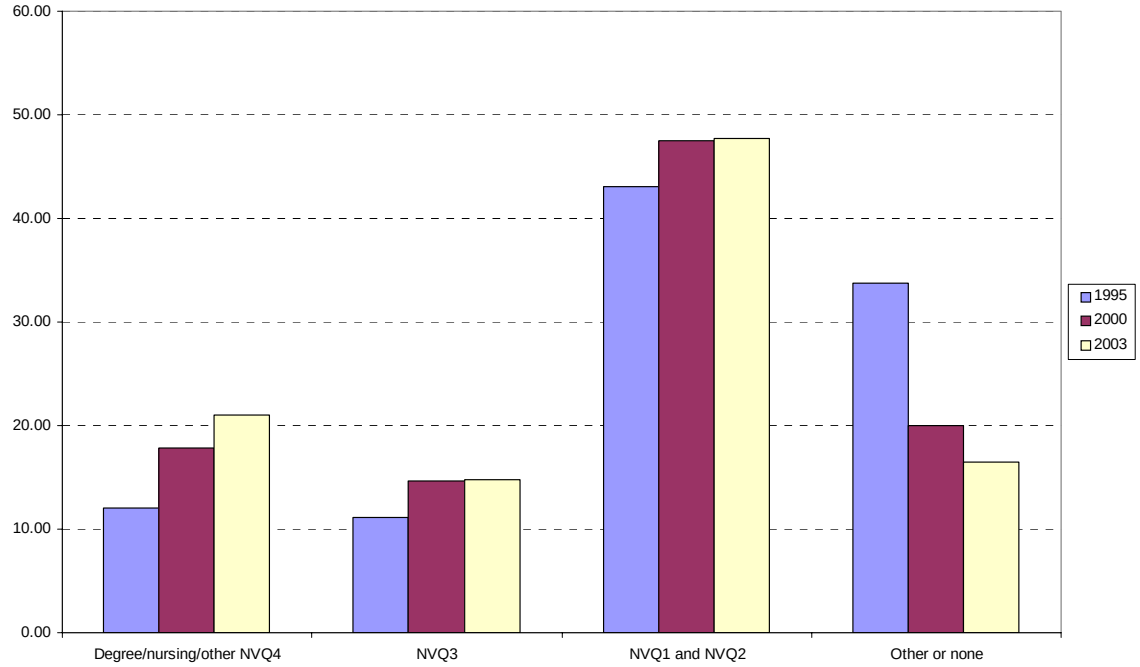
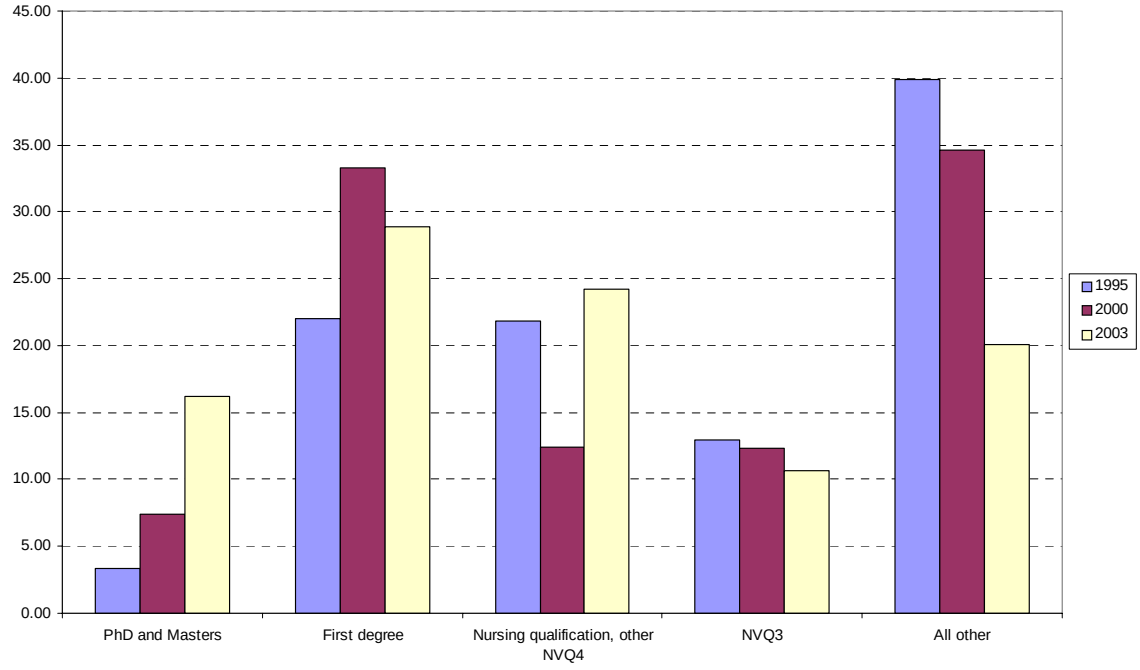


Figure 8.3 Managers: proportions employed by highest qualification held



.

It is also worth considering skill use in the NHS with that for the economy as a whole or other service sectors. Table 8.3 shows skill proportions in the total economy and private market services for the three years shown for the NHS above.

Compared to the total economy, the NHS employs proportionally more workers at the high end of the skill distribution. These differences are more pronounced when the comparisons is between the NHS and private market services. Over this time period growth in proportions of the workforce with the highest qualifications (degree and above) has been higher in the NHS than in the total or market services and the reduction in the use of unskilled workers has been greater.

One interesting trend that can be considered from the comparison between the NHS and other sectors is the extent to which those with nursing qualifications are leaving the NHS to work in other sectors. In 2003 the LFS shows about 550,000 individuals whose highest stated qualification is a nursing qualification. Of these 42% work in NHS Trusts whereas in 1995 nearly 47% worked in NHS Trusts. Some of this attrition can be traced to other areas of the public health sector or private health sectors but much is to other non health industries. Thus 42% of persons whose highest qualification was nursing worked outside the health sector in 1995 and this had risen to 45% by 2003. While this suggests some increase in attrition rates it may well be an underestimate since many people leaving may have other qualifications higher than nursing.

Table 8.3 Skill proportions of the Workforce: NHS selected years

	1995	2000	2003
Total Economy			
Higher degree (Masters, PhDs)	2.7	4.6	5.6
Degree	11.6	13.1	13.8
Nursing qualification/other NVQ4 (higher education below degree)	9.5	9.7	9.9
NVQ3 (A-levels or equivalent)	12.1	16.3	17.8
NVQ1 and NVQ2 (GCSEs or equivalent, vocational qualifications)	40.0	36.5	34.6
Other	7.9	8.0	7.7
None	16.2	11.8	10.6
Market Services[1]			
Higher degree (Masters, PhDs)	1.9	3.1	3.7
Degree	10.9	13.1	13.5
Nursing qualification/other NVQ4 (higher education below degree)	5.9	6.6	6.7
NVQ3 (A-levels or equivalent)	13.8	17.6	19.5
NVQ1 and NVQ2 (GCSEs or equivalent, vocational qualifications)	42.7	38.9	36.7
Other	8.6	8.6	8.5
	16.2	12.1	11.4

1. Comprising distribution, transport, communications, hotels & catering, business services, financial services and personal services.

The changes in the skill use pattern noted above changes aggregate labour input only if there are significant differences in wage rates across skill groups. Table 8.4 shows wages relative to the lowest group for the qualifications given in the Table above. In fact there are very large differences in the wages paid to workers in the NHS with on average those with higher degrees earning about 4 times the average unskilled wages. Note however that these differentials tend to be smaller in the public sector than in the private sector. Also wage differentials have tended to stay reasonably constant through time. Therefore, based on these data, quality change is driven by greater employment growth among skilled workers rather than by any movement to employing more expensive and under competitive assumptions more productive workers, within these skill groups.

Table 8.4 Wages relative to the unskilled, selected years

	1995	2000	2003
Higher degree	3.88	4.07	3.81
Degree	2.72	2.61	2.68
Nursing qualification/other NVQ4	1.88	1.87	1.87
NVQ3	1.39	1.33	1.21
NVQ1 and NVQ2	1.14	1.14	1.09
Other	1.48	1.41	1.70
None	1	1	1

Combining the information on skill use and wages gives growth in quality adjusted labour. Table 8.5 below shows growth in headcount, quality adjusted labour and the percentage point contribution of quality change for all NHS workers, and selected occupation groups.

Table 8.5 Growth in volume and quality adjusted labour input, Total NHS

	Numbers	Quality-adjusted	Difference
Total			
1995-03	**2.68**	**3.47**	**0.79**
2000-03	**4.50**	**5.19**	**0.69**
Doctors			
1995-03	3.30	3.33	0.03
2000-03	4.03	4.20	0.17
Nurses			
1995-03	2.75	3.00	0.25
2000-03	4.24	4.32	0.08
Nursing aux.			
1995-03	2.22	2.34	0.13
2000-03	2.71	3.72	1.01
Health associate profs			
1995-03	3.30	3.39	0.09
2000-03	4.48	4.64	0.16
Healthcare assistants			
1995-03	2.89	3.60	0.71
2000-03	5.00	5.64	0.64
Managers			
1995-03	6.59	8.10	1.51
2000-03	11.18	12.90	1.72
Administrative			
1995-03	2.41	2.93	0.52
2000-03	5.35	6.17	0.82

Again it is useful to compare these trends with other sectors of the economy. In addition we also show growth in labour input for the hospital sector alone, since calculations for other inputs discussed below are carried out mainly for the hospital sector. In this calculation we also include the adjustment for agency workers and convert to an annual hours basis. The sample sizes in the LFS were such that it was not possible to include occupation specific hours or agency adjustments in the previous table. Adjusting for agency workers raises the annual average growth rate over the entire period from 2.68% to 2.78% and since 2000 from 4.5% to 4.63%,

small upward adjustments. Adjusting for trends in hours raises the growth rate further to 4.75% from 2000-2003 but lowers the growth rate marginally over the entire period. Ideally, in the calculations we would like to incorporate trends in hours by qualification group. However when we attempted such a division in the LFS the resulting weekly hours turned out to show implausible variations from year to year and so this calculation was not attempted.

The table shows growth rates over the six years considered in the output calculations. It shows much higher quality adjusted labour input growth in the total NHS and in hospitals than in the economy as a whole. Quality adjustments were similar in percentage points to those in the total economy and larger in the more comparable market services sector proportionally more important in the total economy or market services than in the health sector averaged across the entire time period and the period since 2000. However the aggregate and private sector estimates have been more sensitive to the stage of the business cycle and hence to the starting year chosen. In the past few years the NHS adjustments have been greater.

Table 8.6 Growth in volume and quality adjusted labour input, comparison between the NHS and other sectors, annual average 1999/00-2003/04

	Annual hours	Quality-adjusted	Difference
Total NHS	3.37	4.24	0.87
Hospitals	3.58	4.35	0.77
Total Economy	0.92	1.74	0.82
Market Services	0.71	1.36	0.65

8.2.5 Quality adjustments: refinements

So far the calculations assume that the only variation in type of worker is skill. We also calculated a number of variations which accounted for a further disaggregation of

doctors, regional variations, on the job training and country of birth of workers. The rationale for the first is that since all doctors must have degrees or equivalents in order to practice, the LFS data are not capturing skills acquired through on the job learning. To take account of this we need to disaggregate doctors by type, e.g. consultants, registrars, junior doctors. Data from the Census for the secondary care sector, coupled with snapshots of relative wages from the NHS earnings survey were employed to achieve a crude adjustment based on a division of doctors by consultants and others. Over the period from 2000/01 to 2003/04 this resulted in growth in quality adjusted doctors about 30% above numbers employed. For the same period quality adjusting based on certified qualification led to only an increase in growth of 4%. Therefore at least for doctors, using qualification data alone may not be sufficient to capture all quality change. However since doctors represent less than 1.5% of the NHS Trusts wage bill this amount to only a very small adjustment for the hospital sector. As we do not have comparable data to consider GPs, the adjustment for overall NHS activity is even smaller.

It is likely that workers have received some additional training beyond that associated with their certified qualifications. Using LFS data we can divide workers according to whether they received any job related training or education during the 13 weeks previous to the survey date. Since the data are averaged across quarters this variable picks up any job related training carried out over entire years. For 2003/04 just over 50% of NHS workers answered yes to the question of whether they received some job related training over the past year and the percentage is high across all occupation groups (table 8.7). Nearly half the training occurs in the employees workforce with about 30% off site in education institutions (Table 8.8). About half the workers engaged in training with duration lasting one month or less while about 20% was on going at the time of the survey. Nearly 30% of training was for more than 6 months.

Table 8.7 Percent of NHS workers receiving job related training, 2003/04

Occupational group	%
Total	**50.2**
Managers	56.5
Medical practitioners	68.2
Other health professionals	62.7
Science & engineering professionals, technicians, IT, research	44.0
Teaching & all other professionals etc, plus skilled trades	34.1
Nurses, midwives	64.0
Other health associate professionals, therapists	57.2
Nursing auxiliaries and assistants	48.8
Other healthcare and related personal services	46.2
All other occupations	25.5

Table 8.8 Location of job related training

Location of job related training	%
Employer/another employer's premises	49.1
Training centres etc	11.5
Educational institution	30.9
None of these	8.5

Job related training can vary from one day courses on health and safety to extensive formal learning from senior colleagues. If most training were concentrated in the former then this would have very little impact on overall productivity of workers whereas the latter type of training is likely to have a significant impact. Consistent with the method employed in the remainder of this section, we look at relative wages as an indicator of relative productivity. In 2003/04 workers who gave a positive response to the on the job training question earned on average 28% more than those who gave a negative response. The differential was greatest among highest skilled

groups but was also large across all skill divisions.

In order to account for on the job training we compared the quality adjustments when labour is divided by skill level with those where we included the further division according to whether they received job related training or not. The results of this calculation were to raise the quality adjusted labour input growth rate on average from 1999/00 by about 5%, a small but not insignificant impact. Therefore our estimates of quality adjusted labour were scaled up to reflect the impact of on the job training.

Finally we tried two additional calculations using a division by region and by country of birth but neither significantly altered the results.

8.3 Conclusion

This section considered trends in the use of labour input in the NHS. It showed that labour input growth has been rising very rapidly in recent years, mainly due to growth in the numbers of workers employed but also there is a significant contribution from upskilling of the workforce. The latter is important in understanding why expenditure on the NHS has been increasing rapidly in recent years. Thus a crude calculation suggests some 20% of payments to labour is due to paying for higher skilled workers.

9 Experimental productivity estimates

This section considers measurement of inputs and combining these with output estimates from previous sections to calculate productivity. Lee (2004) reported productivity estimates for the total NHS with details of calculations given in Hemingway (2004). The focus has been on refining the estimates of labour input to take account of the use of various types of skilled labour. Labour is by far the most important input used in producing health services, accounting for about 75% of total hospital expenditures. These measures can then be combined with the aggregate and hospital output measures given in Section 7 to calculate labour productivity growth

rates. Total factor productivity (TFP) estimates are then calculated for the hospital sector and the NHS as a whole. In the hospital sector it is straightforward to combine the direct measures of labour input discussed in the next section with data on the use of intermediate inputs and capital from Trust Financial Returns to calculate TFP growth. ONS data are combined with the estimates for labour input from Section 8 to derive the aggregate TFP estimates.

9.1 Labour input and labour productivity growth

Using the calculations reported in section 8, annual estimates of the volume of labour input, and quality adjusted labour input, are shown in the following table for the aggregate NHS and the Hospital NHS.

Note the hospital sector captures activities that are recorded in HES, which were the focus of Chapter 5, *and* non-HES activities such as outpatient, A&E, rehabilitation and mental health activities. This broader definition of outputs corresponds to the basis on which hospital inputs are measured. The overall NHS includes activities such as primary care, prescriptions, community care, dental and ophthalmic.

Table 9.1 Volume of labour input and quality adjusted labour input, annual estimates of growth rates

	Labour input: volume		Labour input: quality adjusted	
	Total NHS	Hospital	Total NHS	Hospital
1998/99-1999/00	1.58	1.49	2.53	3.47
1999/00-2000/01	1.05	1.77	1.45	2.75
2000/01-2001/02	5.42	5.05	5.31	3.77
2001/02-2002/03	4.69	4.51	5.57	5.91
2002/03-2003/04	4.48	5.47	4.94	5.83
Average	3.43	3.64	3.95	4.34

Labour productivity growth rates are derived by taking the growth in output minus the growth in labour input. Here we show calculations for unadjusted cost weighted output and the quality variants chosen for illustrative purposes in Section 7 adjusting for survival and waiting times. Denote these two variants by Q1 and Q2. For comparison purposes these output growth rates are shown in Table 9.2.

The first panel in Table 9.3 presents labour productivity estimates based on volume of labour and shows positive average labour productivity growth across the period for all output variants for the total NHS and for the Q1 variant for hospital output but a small negative number for unadjusted hospital output. When quality adjusted labour is used instead, average labour productivity growth becomes negative in the total NHS when output is not quality adjusted, but is approximately zero using the Q1 variant of quality adjusted output. Hospital labour productivity growth is negative for all output variants when quality adjusted labour is used. Note negative labour productivity growth in health services is not unusual in international comparisons. Data from the US national accounts suggests labour productivity growth, unadjusted for labour quality changes, was -0.31% on average from 1999 to 2002. Adjusting for labour quality would reduce this further.

Table 9.2 Output growth

	Total NHS				Hospital		
	Unadjusted	Q1	Q2		Unadjusted	Q1	Q2
1998/99-1999/00	2.61	1.96	2.22		2.03	0.79	1.28
1999/00-2000/01	2.11	2.46	2.26		1.54	2.16	1.80
2000/01-2001/02	3.85	3.82	3.74		4.48	4.43	4.31
2001/02-2002/03	5.07	6.11	5.78		3.94	5.57	5.06
2002/03-2003/04	4.43	5.20	4.93		4.78	6.00	5.56
Average	3.62	3.91	3.79		3.35	3.79	3.60

Notes: Q1 is the 'high' quality adjustment variant with $k = q^0/q^* = 0.8$ if $a - k > 0.05$, and $k = 0$ otherwise, discounts to date of treatment with charge for wait, discount rates on waits and life expectancy equal to 1.5% and where the waiting time variable is the 80th percentile wait in each HRG. Q2 is our recommended quality variant and sets $k = q^0/q^* = 0.8$ for electives, $k = 0.4$ for non-electives, k = actual k for those HRGs where known, if $a - k > 0.10$, and $k = 0$ otherwise; discounts to date of treatment with charge for waits, discount rates on waits and life expectancy equal to 1.5% and uses 80th percentile waits.

Table 9.3 Labour productivity growth

	Total NHS				Hospital		
	Unadjusted	Q1	Q2		Unadjusted	Q1	Q2
Labour volume							
1998/99-1999/00	1.02	0.38	0.63		0.53	-0.69	-0.20
1999/00-2000/01	1.05	1.40	1.20		-0.23	0.38	0.03
2000/01-2001/02	-1.49	-1.51	-1.59		-0.54	-0.58	-0.70
2001/02-2002/03	0.35	1.35	1.04		-0.55	1.02	0.53
2002/03-2003/04	-0.05	0.69	0.42		-0.65	0.51	0.09
Average	0.17	0.46	0.34		-0.29	0.13	-0.05
Quality adjusted labour							
1998/99-1999/00	0.08	-0.56	-0.30		-1.40	-2.60	-2.12
1999/00-2000/01	0.65	1.00	0.79		-1.18	-0.57	-0.92
2000/01-2001/02	-1.38	-1.41	-1.48		0.68	0.64	0.52
2001/02-2002/03	-0.48	0.51	0.20		-1.87	-0.32	-0.81
2002/03-2003/04	-0.48	0.26	-0.01		-0.99	0.16	-0.25
Average	-0.32	-0.04	-0.16		-0.95	-0.54	-0.72

Note: all productivity estimates use geometric means

9.2 Intermediate and capital inputs

Intermediate input for the hospital sector comes from the Trust Financial Returns (TFR) and is deflated by a modified version of the DH Health Services Cost Index (HSCI) to derive a volume measure. Intermediate input was defined as all current non pay expenditure items in the TFR, and hence excluded all purchases of capital equipment and capital maintenance expenditures as these items cannot be allocated to a particular year's output.

The list of items included and their shares in total intermediate expenditure in selected

years is shown in Table 9.4. This shows the share of drugs increasing rapidly and a declining trend in the miscellaneous category with no other intermediate category showing much change. Within the final category, external purchase of health care from non-NHS bodies has shown an increased share through time but remains small at about 6% of total intermediate in 2003/04.

Table 9.4 Share of intermediate expenditure by type

	1998/99	1999/00	2000/01	2001/02	2002/03	2003/04
Drugs	0.238	0.242	0.286	0.303	0.316	0.338
Other Clinical Supplies	0.041	0.042	0.053	0.055	0.052	0.056
General Supplies and Services[1]	0.139	0.134	0.150	0.144	0.128	0.122
Establishment expenditure[2]	0.166	0.159	0.187	0.176	0.151	0.143
Non-capital premises[3]	0.112	0.101	0.115	0.114	0.101	0.104
Other[4]	0.304	0.322	0.209	0.209	0.252	0.237

1. Hotel, catering and cleaning services; 2. Stationery, communications, advertising and transport costs; 3. Energy, rent and external services; 4. Miscellaneous services including external purchase of health care from non-NHS bodies and other external contract expenditures.

These numbers for intermediate input were deflated by an aggregate price index, derived as a chain linked index of corresponding HSCI items. This resulted in a very small upward adjustment in the intermediate input deflator than one using all items in the HSCI, as the prices of capital items have been growing more slowly than current items and in the case of computers have been falling.

To be consistent with the methodology employed by ONS, capital input could be measured by depreciation from the TFR. However this calculation ignores any capital services from capital purchases in the current year. An alternative is to assume a proportion of these expenditures are depreciated in the current year. Since much of the expenditure is on computers, software and medical equipment with low asset lives, we assume one third of these assets are depreciated. These capital services are deflated by a chain linked deflator for capital items in the HSCI while depreciation is deflated by

the ONS capital consumption deflator for the total NHS. Adding a proportion of capital expenditures implies average annual real capital growth of 4.76% per annum from 1998/99 to 2003/04 compared to 4.69% using depreciation alone, a small effect. However it does raise capital's share from 0.041 to 0.080. Finally consistent with the methodology employed in private services the value of business rates are added to capital's share.

Calculating input shares is more difficult for the total NHS since we attempt to combine data from different sources. For the aggregate NHS we use TFR data on payments to labour in the hospital sector and use ONS data for payments to labour in other parts of the NHS. Similarly, intermediate inputs are derived combining expenditures from TFR and PFR with ONS data on other parts of the NHS. Capital inputs are those employed by ONS in their measures of Health Sector Productivity. Family Health Drugs are deflated by the cost of all items rather than the ONS quality adjusted Paasche variant. The estimates for the NHS should be treated with considerable caution since data are being taken from a number of sources which may need further reconciling.

Table 9.5 shows average period input shares and average growth in the three inputs where labour input is the quality adjusted variant.

Table 9.5 Average period input shares and average growth in inputs

	NHS		Hospital	
	Shares	Input growth	Shares	input growth
Labour	0.61	3.95	0.72	4.34
Intermediate	0.33	8.58	0.20	4.93
Capital	0.06	3.12	0.08	4.76

Labour represents a lower share in the total NHS than in the hospital sector, mainly due to the inclusion of family health prescribing in the former. Growth in intermediate input is very large in the total NHS, while all three inputs show similar average growth rates in the hospital sector.

9.3 Total factor productivity growth

Combining the input shares with growth in real inputs allow the calculation of total input growth and subtracting this from output growth yields total factor productivity growth rates, shown in Table 9.6. Average TFP growth rates are strongly negative for the total NHS for all quality variants. The numbers based on quality adjusted output are similar to those calculated by ONS reported in Lee (2004). The ONS estimates using the most comparable methods to measure inputs suggest average TFP declining by 1.34% per annum over the same period. However it should be noted that ONS output measures use reference cost activities which are not directly comparable with the HES based data employed in this report's calculations. Average TFP growth is also negative for the hospital sector but less so than for the total NHS. The results are sensitive to the allocation of expenditures between the three broad categories of inputs. Lee (2004) highlighted the sensitivity of the results to the deflators used, in particular for drugs – see also the discussion in section 10.5. However since drugs are both an input and an output this sensitivity is surprising. Hence some further investigation is required and will be carried out following discussions with ONS.

Table 9.6 Total factor productivity growth

	Total NHS				Hospital		
	Unadjusted	Q1	Q2		Unadjusted	Q1	Q2
1998/99-1999/00	-2.33	-2.95	-2.71		-2.82	-4.00	-3.53
1999/00-2000/01	0.55	0.89	0.69		0.30	0.91	0.56
2000/01-2001/02	-2.12	-2.15	-2.22		0.17	0.13	0.01
2001/02-2002/03	-1.86	-0.88	-1.19		-2.01	-0.46	-0.95
2002/03-2003/04	-2.97	-2.25	-2.51		-1.13	0.02	-0.39
Average	-1.75	-1.48	-1.59		-1.11	-0.70	-0.87

The finding that TFP growth is negative is not unusual in the private sector. For example Basu *et al*. (2003) report negative gross output based annual average TFP

growth rates for a number of sectors in the 1990s including insurance and business services. Similar results have frequently been reported by the US Bureau of Labour Statistics. Negative TFP growth is mostly likely to occur in service sectors where output is poorly measured and quality adjustment is minimal. TFP growth rates for the private sector using comparable measurement methods are not yet available for the period under consideration in this report.

When inputs are measured correctly, with adjustments for quality change then the TFP residual is close to a measure of pure technical change so long as output is also measured correctly. But as emphasised in many parts of this report, we are only capturing part of the improvement in quality of care via our proposed adjustments for survival, health effects and waiting times. Because of this incomplete adjustment for quality change we expect to underestimate TFP growth. There are also reasons why in the short term at least we might expect negative growth rates. The literature on the impact of information technology on productivity in the private sector points to an important role of organisational changes in facilitating benefits from new technology. Basu *et al*. (2003) suggest that these changes can lead to declining TFP in the short run due to disruption of production processes. There is no doubt that the NHS is undergoing significant change.

Of more consequence for the health sector is the notion that there are diminishing returns as increased activity allows treatment of more complex and hence most costly cases. Activity rates have been increasing more rapidly in recent years. Some evidence in support of this is provided by the increased average age of patients treated in hospitals, from 48.6 years in 1999/00 to 50 years in 2003/04. In addition there has been some increase in the expenditure shares of HRG categories with the title 'complex elderly' from 3.4% of expenditures to 4.2% over the same period. Changes in the case mix are likely to be larger within than across HRGs but we lack the necessary data to examine this. Data that identified the characteristics of patients would also be useful in identifying the extent to which changes in NHS productivity are affected by diminishing returns.

10 Improving the data

We appreciate that the Department of Health is trying to reduce the amount of data collected from the NHS (Review of Central Returns Unit). However, the data requirements outlined below are essential to development of robust measures of output, productivity and quality change in the NHS. These data will be required not only by the DH and Trusts to improve performance of the NHS but also by outside bodies such as the Treasury and ONS.

10.1 Outcomes data

10.1.1 Health outcomes

The main aim of the health system is the improvement of the health of the population. This being so, it would seem reasonable that any measure of health system performance, output and productivity should include measures of the effect of the system on health. The challenges associated with measuring the effect of interventions are discussed in Appendix C.

The construction of a productivity index requires information about changes in health status attributable to interventions. Such information currently is not collected by the NHS. We *suggest* the systematic use of a standardised measure of health status to improve the effective management of the NHS and to provide the fundamental data needed to properly reflect changes in NHS productivity.

We have suggested that the NHS should collect data on the health of patients before and after treatment. An outcome measure based on the difference between snapshot measures of health status before treatment h^b and after treatment h^a is an imperfect measure of the change in the discounted sum of QALYs due to treatment. It does not measure health with and without care but health before and after care. It also replaces each time profile with a single snapshot. For some treatments and conditions the effect of treatment is merely to slow down the rate of decline in health status, so that $h^a - h^b < 0$ even though the treatment increases the sum of QALYs compared with no

treatment. For these cases incorporating estimates of h^b and h^a in estimates of the *level* of output in any given year would reduce measured output.

However, since the aim is to measure the *rate of growth* of output productivity we are interested in whether the rate of growth of $\Delta h = h^a - h^b$ is a reasonable approximation to the rate of growth of the effect of treatment on the discounted sum of QALYs. The important issue is how well the rate of change in measures based on the snapshots h^b, h^a approximates the rate of change in the areas under the two time profiles of health streams with treatment $h^*(s)$ and without treatment $h^o(s)$

Both the level of health before treatment h^b and the health of treated patients if not treated depend on the patient population selected for treatment and on the general health of the population. It is not unreasonable to suggest that the rates of change of h^b and the discounted value of the no treatment health profile $h^o(s)$ over time will be similar. Both the snapshot level of health after treatment h^a and the discounted value of the time profile h^* will be measured on the same population and hence are affected by the same factors including any technological change.

Hence, despite the imperfections of the difference between snapshots of post and pre treatment heath status for calculating the *level* of productivity, we suggest that rates of change of measures based on h^b, h^a will improve estimates of NHS output growth compared to estimates where such information is not used.

The following points need to be addressed in such a data collection exercise.

- **NHS patient sample**. Since the scale of NHS activity is so broad and the potential volume of patients is so large sampling seems a more sensible strategy than attempting to measure health effects for all NHS patients in a sector. Sample sizes are likely to vary across different types of patients. While a random sample of the NHS patient population would be preferable, in the first instance it may be satisfactory to undertake a pilot exercise at a handful of Trusts. This might be adequate for national measures of output but would be inadequate if information on health outcomes are to be used to improve performance of the NHS.

- **Choice of instrument**. A single generic, rather than condition-specific instrument is required in order to facilitate aggregation across different types of NHS activity. Profiles such as SF-36 provide multiple measures of outcome but are unsuitable for most non-clinical purposes since they typically lack the capacity to form a single aggregate index. Derivatives of SF-36 such as the SF-6D do not suffer from this deficiency. The EQ-5D is designed to produce a single index and its five dimensions have been calibrated in terms of social preference weights of a UK population and is probably the primary candidate measure.

- **Timing**. The timing of before and after health status measurement may depend on the type of activity (emergency or elective) and on diagnostic category or intervention type since different treatments may have an effect over shorter or longer periods. Our analysis of data for two elective procedures (hip and knee replacement) shows most treatment gain after six months but clinical advice could be used to determine the appropriate period for follow-up post treatment for other conditions.

- **Grouping of NHS activities**. Given the enormous range of NHS activities it is necessary to group them for data analysis. The main grouping of secondary care activities is at present by HRG which attempts to group activities by their costs. But a given HRG may contain a large number of procedures which have very different effects on health. The availability of patient-level health outcomes data by ICD will permit matching to other datasets (such as HES) and will make it possible to explore the extent to which health outcomes are related to other routinely collected patient characteristics, such as age, gender, diagnoses, procedure, survival rates and mortality. This is essential if we are to develop disease specific studies of improvement in health care.

- **Frequency of data collection.** If the pace of technological change in medicine was slow enough it could be argued that collecting data on health outcomes was an exercise that needed to be undertaken only at intervals of several years. But technological change in medicine and pharmaceuticals is

rapid, the NHS is subject to frequent organisational change which may affect the mix of patients receiving particular treatments and the speed at which new technology spreads. We believe that only a continuous sampling of the NHS patient population will be adequate to capture trends in the impact of NHS services on patients.

10.1.2 Feasibility

Outside clinical trials, experience of routine collection of health status data in the UK is patchy. Individual clinicians and clinical teams make use of a variety of standardised measures, but this is largely uncoordinated, its coverage remains undocumented and aggregation of such data is problematic given the use of different instruments.

Although there are a limited number of examples of prospective health data collection these examples demonstrate that such data collection would be feasible in the NHS.

- The survey of acute inpatients conducted by **Picker International** showed that it was possible to collect EQ-5D data from a sample of patients recently discharged from all NHS Trust hospitals. The value of these data would have been enhanced if they been linked with basic HES variables, such as diagnosis or procedure. Given current DH policy to regularly survey patient experience, inclusion of questions on health outcomes would involve very low marginal cost.

- The **Health Outcomes Data Repository (HODaR)** operates a continuous survey of all inpatients and outpatients at a single large Welsh Trust. These are now linked to individual level primary and community care data. Data for more than 30,000 patients have been collected, almost 10% of these having completed EQ-5D on more than one occasion. However, the data are predominantly based on post-discharge observations and this limits their value in measuring health outcomes. Since the advent of this project the HODaR survey has started to collect data on pre-admission health status.

- As described in Appendix C, for a number of years **BUPA** has been routinely administering health status questionnaires to patients before and three months after treatment, with some 100,000 having now been surveyed. These data are restricted primarily to elective procedures. BUPA plan to extend them to four types of cancer.

10.1.3 Cost

The incremental costs of introducing systematic observation of health status via existing information systems are difficult to estimate. Currently BUPA estimates that it costs around £4 per patient to administer their manual system of health status measurement based on SF-36.

The introduction of systematic health status measurement might be achieved under the aegis of the National Programme for Information Technology announced in December 2002 with a budget of £2.3 billion and which the Audit Commission suggested would provide the Department of Health with the opportunity to improve NHS data quality.

It would seem sensible to consider an extension to the current HES-based data to provide maximum scope for exploitation through record linkage. Modification of this sort ought not to incur a significant cost. However, the data captured from patients will require additional organisational and administrative costs. Patient-centred reporting systems using traditional paper and pencil techniques require costly processing in order to link them to other NHS data. Computer-assisted interview methods have scope for more efficient data acquisition and transmission but would need more costly administration. The use of handheld PDA recording systems is now becoming a feature of many clinical trials that record patient-reported health status and it can be expected that hardware costs will continue to fall.

10.2 Other outcome measures: patient satisfaction

As we emphasised in our First Interim Report the effect of the NHS on health status is

obviously an extremely important dimension of NHS outcomes but other dimensions, especially process related outcomes, should also be taken into account. However, existing patient surveys rarely ask similar questions over time. The DH should decide on core aspects of the patient experience and ensure these questions appear each year in patient satisfaction surveys. A problem we discuss in this report is that of identifying a relevant weight to place on patient experience in a quality adjusted output index. Ryan (2004) recommended that the DH undertake a series of discrete choice experiments to obtain evidence on the relative value patients attach to different aspects of process quality. We endorse this recommendation.

10.3 General practice data

10.3.1 GP activity

We have discussed the measures of GP consultations in section 4.4, noting the unsatisfactory nature of the current source (General Household Survey) and that the DH is now planning to obtain consultations directly from GP record systems via the QRESEARCH database. We have agreed to investigate this new data for the DH to determine what if any adjustments need to be made to improve its reliability.

Ideally NHS productivity measures should be based on numbers of patient journeys of different types where journeys are likely to involve both primary and secondary care. In the absence of routine record linkage such measures are not currently feasible but it would still be worthwhile getting a finer breakdown of GP consultations to allow for the changing mix of providers and for the changing mix of types of consultations.

We have also discussed measures of general practice quality in section 4.13. Whilst QRESEARCH and similar databases of GP record systems and the central collection of the greatly enlarged set of quality indicators linked to GP pay will provide potentially useful quality adjustment these will need to be based on careful empirical modelling of GPs responses to the new financial incentives, especially possible diversion of effort from unremunerated quality efforts.

10.3.2 GP cost weights

The PSSRU estimates the unit costs of GP and nurse consultations (http://www.pssru.ac.uk/pdf/uc2003/uc2003.pdf) using a variety of official and unofficial sources. Several of the estimates rest on self reported GP activity from the 1992/3 GP Workload Survey undertaken for the DDRB. There does not appear to be a more recent survey of GP activity and we recommend that DH should consider undertaking such a survey at regular intervals.

10.3.3 General practice staff

Prior to April 2004, practice staff such as practice nurses, although practice employees, were partly paid for by the NHS and so a record was kept, though it was not a reliable source because not all practices claimed these subsidies. Under the new GP contract the subsidies have been abolished and there is no record of non-GP staff in practices. Now practices instead receive a sum of money based on their practice. We recommend that the DH make it a condition of the practice contract that a full return of employed staff is made.

10.3.4 Prescribing

The prescription activity measure in the recently revised NHS outputs index is derived from PPA data. The PPA data are collected in order to remunerate pharmacists (and dispensing GPs). It is therefore a comprehensive measure of prescriptions dispensed and can be disaggregated to product type if required. The data are reliable, comprehensive and readily available at national levels of aggregation. They have been used to construct a number of indicators of practice prescribing quality as well as quantity.

The usefulness of the data could be greatly improved and this would be relatively simple. The most obvious example is by improving the patient information on the prescription form. At the moment the only patient data on the form indicates if the patient is entitled to free prescriptions and on what grounds. The information has been used by the Prescribing Support Unit to produce the Low Income Scheme Index (LISI) which measures the proportion of prescriptions which are dispensed without

charge on grounds of low income. The LISI is the only direct variable measuring practice population socioeconomic status which relates directly to practice patients rather than being attributed from Census or Social Security data on the basis of patient postcode. Adding a field for diagnosis to the prescription form would greatly enhance the usefulness of routine prescribing data as a measure of prescribing quality. Adding gender and age fields would also improve the socioeconomic data and improve prescribing quality indicators. We recommend that the DH should add these fields to the prescription form.

10.4 Other primary care data

NHS Direct, NHS Direct Online and Walk-In Centres are recent innovations in the provision of first contact advice and information. They are likely to reduce the costs to patients of such first contacts, leading both to an increase in primary care activities and to a change in the mix of activities in general practice. The organisations are expected to play an increasing role in the NHS over the coming years and it is important that their presence is recognised in measures of NHS output and productivity.

Aggregate data on use of NHS Direct and NHS Direct Online are available. In order to measure the outputs of the services more accurately it would be helpful to have data on

- the breakdown of enquires between the provision of health advice and information about the health service
- the type of conditions people seek health advice about
- actions that are recommended as a result of the request

It is possible that such data have been collected, for example via the website service for those who seek advice from a nurse which involves self-completion of a detailed questionnaire on the nature of the symptoms and condition, as well as personal information. Presumably – although we have not been able to ascertain whether this is the case – enquiries that result in a self-care recommendation are logged also. The telephone service seems to be set up in a similar fashion, the difference being that the

information is recorded by NHS Direct staff.

It appears, therefore, that these two organisations routinely collect (or, at least, have the capacity to collect) detailed electronic information from every person making an enquiry about their (or their family member's) health condition.

We recommend that the DH should utilise more of the data collected by NHS Direct to improve measures of output.

10.5 Inputs

It is important to have a comprehensive coverage of inputs used in producing health services in order to explain changes in outputs and to measure productivity. Thus we require values of expenditures on inputs, volume measures and price deflators to convert values to volume measures when the latter are not available. We also require data on the extent to which the quality of inputs are changing through time. When volume measures are available, under the assumption that payments to labour equal marginal products, weighting diverse inputs by their shares in wage bills can be employed to adjust for quality change. Alternatively hedonic regressions may be employed to quality adjust price deflators. Both methods depend on competitive market assumptions which are unlikely to characterise many of the markets in which the NHS operates.

The analysis of labour input showed that it is possible to derive reasonable volume measures and adjustments for quality change by linking readily available data sources. The quality adjustment employed was certified qualifications with an additional adjustment for job related training. While certified qualifications are useful they may not be sufficiently detailed to capture differences in the productive capacity of some employees. Many professionals within the NHS have similar qualifications since there are minimum requirements set by professional bodies. While the NHS census does include very detailed data on numbers of professionals by grade, there is no comparable data on earnings. The persons responsible for the Earnings survey within DH were unwilling to attempt to match earnings to the Census numbers by type on the

valid grounds that the sample size was too small. A larger sample survey of wage rates and earnings within the NHS would be very useful, not only for the productivity calculations carried out in this report but also to calculate the extent to which increases in payments to labour are due to employing more highly skilled personnel as against mere wage inflation.

The productivity calculations in section 9 above highlight the importance of intermediate inputs in overall NHS activity and to a lesser extent in hospital activity. Reliable volume measures require reasonable estimates of intermediate input deflators. The main problem with intermediate inputs is the price deflator employed for drugs. As mentioned in section 3.3 above ONS is currently attempting to refine its deflator for prescription drugs based on detailed data on prices. Data sources for prices of hospital drugs are not readily available. Even if such data could be collected, it is doubtful if they would be useful as a tool for quality adjustments. Hedonic type adjustments are only valid in competitive markets. The market for hospital drugs in Britain is best characterised as a bilateral monopoly with a monopsony purchaser buying from powerful oligopoly drug producers. The standard textbook model of bilateral monopoly shows that the price will depend on the relative bargaining power of the purchasers and suppliers. It is well known that the NHS buys drugs at a discount and it may well be the case that discounting on new drugs may swamp any quality change, at least at the point of entry. The use of prescription drugs are in the control of independent GPs so the monopsony element is less important but there remain market imperfections on the producer side.

These remarks suggest that using drugs prices to measure quality requires an understanding of how markets operate and should be based on a time profile of prices rather than comparisons between old and new drugs at the time new drugs enter the market. A more fruitful but also costly approach might be to examine patient outcomes and drug use from disease registers or to obtain opinions from panels of experts. This is simply another example of the need for outcomes data if we are to measure technical change and productivity in health care.

In order to quality adjust capital input it would be useful to have separate investment data on types of equipment that have seen rapid technological change and change in unit cost (e.g. MRI scanners). Alternatively it would be useful to have the value of the stock of these assets, numbers of items and age profiles of the stock. One problem that must be addressed is the gap in data created by PFI confidentiality. We understand that a significant proportion of new investment in equipment such as scanners is being undertaken under PFI contracts. Unless it is possible to access information on stocks and value, it may not be possible to adequately deal with questions of productivity growth and technical change associated with investment in new equipment. It would also be useful to have information on investment by GPs.

11 Conclusions and recommendations

11.1 Methods

11.1.1 The preferred approach

Economic theory suggests that the preferred way of measuring NHS output is with a value weighted output index.

- The unit of output is the patient treated, the characteristics of output valued by individuals indicate quality and the weight attached to each characteristic reflects the marginal social value of the characteristic.
- The index overcomes the serious problem of a cost weighted index where movement to more cost-effective ways of treating patients appears as a reduction in output.
- Data necessary to estimate this index are not currently available but are feasible to collect.
- A condition specific value weighted index can be constructed as data on major diseases becomes available.

Not only is the value weighted index theoretically correct, it would allow measurement of improvements in delivery of services intended to raise both productivity and patient satisfaction.

11.1.2 Methods using existing data

It is not possible to calculate a value weighted index with current data. It is possible to quality adjust the hospital component of a cost weighted NHS output index using existing data combined with some assumptions. We have

- spelt out the methods for quality adjustment with existing data in some detail, taking care to emphasise the necessary assumptions and their implications, rather than merely presenting plausible ad hoc adjustments which may in fact contain dubious assumptions or value judgements.
- shown how it is possible to use routine data on short term survival and waiting times, coupled possibly with an explicit assumption about the proportionate effect of treatment on health, to calculate quality adjusted cost weighted output indices.
- presented experimental calculations of these indices to compare their effects on a simple cost weighted output index and to investigate the empirical implications of the assumptions about important parameters which are required.
- used data on the health effects of a limited set of treatments to illustrate the construction of a value weighted index for the set and to shed further light on the implications of making possibly inaccurate assumptions about the health effects when constructing an index for all hospital treatments.
- described how it is possible in principle to use data on other aspects of care (readmissions, MRSA, patient satisfaction with food, cleanliness and non-clinical care) to provide an additional quality adjustment and have produced some illustrative examples of such adjustments based on the current unsatisfactory data on these characteristics of care.
- described a method of quality adjustment using the information on longer term survival which will become available in the near future.

11.2 Results

11.2.1 Results for the hospital sector

Results are reported for a cost weighted quality adjusted output index over the period 1999/00 – 2003/04. It was only possible to quality adjust for hospital activity, 47% of expenditure covered in the DH cost weighted output index. In Section 5 we examine sensitivity to discount rates and key assumptions. Table 11.1 summarises the central results.

Table 11.1 Hospital sector cost weighted output index with recommended quality adjustments (growth rates%)

	No adjustment	Survival and health effects adjustment[1]	With survival. health effects, and waiting time adjustments[2]
1998/99-1999/00	2.03	0.91	1.28
1999/00-2000/01	1.54	1.99	1.80
2000/01-2001/02	4.48	4.40	4.31
2001/02-2002/03	3.94	5.19	5.06
2002/03-2003/04	4.78	5.51	5.56
Average p.a.	3.35	3.60	3.60

[1] This sets $k = q^0/q^* = 0.8$ for electives, $k = 0.4$ for non-electives, k = actual k for those HRGs included in the specimen index where this is known, provided $a - k > 0.10$ and $k = 0$ otherwise.
[2] Recommended quality variant 2. As note 1 plus discounts to date of treatment with charge for wait, with discount rates on waits and life expectancy equal to 1.5% and uses 80th percentile waits.

Table 11.2 shows the impact of incorporating quality adjustments for the hospital sector into the overall NHS cost weighted output index.

Table 11.2 Aggregate NHS cost weighted output index with hospital sector quality adjustments (growth rates %)

	Unadjusted CWOI	CWOI with hospital survival and waiting time adjustments[1]
1998/99-1999/00	2.61	2.22
1999/00-2000/01	2.11	2.26
2000/01-2001/02	3.85	3.74
2001/02-2002/03	5.07	5.78
2002/03-2003/04	4.43	4.93
Average p.a.	3.62	3.79

[1] Recommended quality variant 2. This sets $k = q^0/q^* = 0.8$ for electives, $k = 0.4$ for non-electives, k = actual k for those HRGs included in the specimen index where this is known, provided $a - k > 0.10$ and $k = 0$ otherwise; discounts to date of treatment with charge for wait, with discount rates on waits and life expectancy equal to 1.5% and uses 80th percentile waits.

Overall, our results show that quality adjustment with existing data can make an impact on measures of NHS output and as more routinely collected data becomes available, the quality adjustment can be improved.

11.2.2 Other quality indicators

We were asked to explore the use of indicators such as patient satisfaction, readmission rates, clinical errors and incidence of MRSA. The data are not at present suitable for inclusion in an output index but some "illustrative" adjustments are reported in Table 11.3.

Table 11.3 Illustrative calculations of hospital CWOI with adjustments for survival, waiting times, patient satisfaction as measured in patient surveys, readmissions and MRSA. Average annual growth rates 2001/2 to 2003/4

Average Growth Rates 2001/2 to 2003/4	% p.a.
Unadjusted cost weighted output index	4.34%
With adjustment for survival and waiting times[1]	5.74%
With adjustment for adjustment for satisfaction with food, cleanliness, and non clinical care[2]	5.71%
With adjustment for MRSA, readmissions satisfaction with food, cleanliness, and non clinical care[2]	5.71%

[1] Variant 1 (k = 0.8, mortality cut off 0.15, discounting to date of treatment, discount rate on waits 1.5%, on life expectancy 1.5%, 80th percentile waiting time measure).
[2] 5% weight on non-clinical care

11.3 Total factor productivity growth

Employing a new quality adjusted index of labour input in the hospital sector, provisional estimates of productivity growth are reported. Key results appear in Table 11.4. They highlight the importance of quality adjustments to output in evaluating NHS performance.

Table 11.4 Total factor productivity growth (%)

	Total NHS			Hospital	
	Unadjusted	Recommended Quality Variant		Unadjusted	Recommended Quality Variant
1998/99-1999/00	-2.33	-2.71		-2.82	-3.53
1999/00-2000/01	0.55	0.69		0.30	0.56
2000/01-2001/02	-2.12	-2.22		0.17	0.01
2001/02-2002/03	-1.86	-1.19		-2.01	-0.95
2002/03-2003/04	-2.97	-2.51		-1.13	-0.39
Average p.a.	-1.75	-1.59		-1.11	-0.87

11.4 Recommendations

Recommendations for improving quality adjustment were made throughout the report. We summarise the main ones here.

For the medium term improvement of the output index improvements to the data are required. We recommend:

- Routine collection of outcomes data for a range of NHS treatments. The programme should start with a few high volume elective and medical conditions that would permit sampling rather than complete coverage. The data would also be immensely useful for other purposes including monitoring of Trust performance and improved cost-effectiveness analysis of particular treatments.

- Collection of longer term survival data by linkage of HES and ONS records to produce estimates of patient life expectancy.
- A patient identifier that will permit grouping NHS activities across institutions and by disease. The DH has plans to implement this change.
- Stated preference studies of patients to establish their relative valuations of the characteristics of NHS output from waiting times to being treated with courtesy and dignity by staff. The studies should also include a cost characteristic so that monetary valuations can be inferred and all characteristics can be valued in a common unit. The studies will enable the data from patient satisfaction studies to be utilised for quality adjustment as well as informing decision making in the NHS.

For the short run improvement of the output index with available data:

- We recommend the use of short term survival coupled with life expectancy to quality adjust hospital output.
- The short term survival adjustment will underestimate output growth. We recommend that it be coupled with an estimated health effect derived from an estimate of the proportionate effect of treatment:
  - As the data become available from surveys of patient health before and after treatment and elsewhere, treatment specific estimates of $k_j = q^o_{jt} / q^*_{jt}$ should be used.
  - Where there are no treatment specific estimates, k_j should be estimated as the volume-weighted mean of existing treatment specific estimates for the relevant class (electives and non-electives).
  - In the absence of any estimates of treatment specific k for non-electives the estimate for non-electives should be equal to half the volume-weighted mean k of the electives.
  - The health effects adjustment should be used only for treatments with a mortality rate of 0.10 or less.
- We recommend the use of a waiting time adjustment based on discounting to date of treatment, with a charge for waiting. Theoretical considerations

suggest that the discount rate on waits should be the same as the discount rate on QALYs. We suggest 1.5%.

- We do not recommend quality adjustments based on patient satisfaction with food, cleanliness, and non-clinical care until there are data on the relative marginal values of these outcomes. If it is felt that estimates of the costs of cleaning and food derived from Trust accounts reasonably reflect marginal social values then it would be possible to include an adjustment just for these satisfaction indicators.
- We do not recommend quality adjustments based on readmission rates and MRSA because of data problems and because they may reflect aspects of care which are better captured in the other quality adjustments.
- We recommend the use of 30 day mortality, rather than in hospital mortality, as the measure of short term survival, since we believe its greater theoretical merits outweigh the difficulties in calculating it. As data linkage methods are improved the advantage of the 30 day mortality will increase.
- The waiting time measure should be a certainty equivalent wait, to avoid the need to calculate adjustments on individual data. The 80th percentile wait seems a reasonable value.
- Quality adjustments of hospital output should use CIPS rather than FCEs as the unit of output.
- HES rather than the Reference Cost data base should be the source of data on hospital outputs.

11.5 Acknowledgements

The outputs and productivity project is funded by the Department of Health. The views expressed here are not necessarily those of the Department of Health. This report has benefited greatly from discussion with numerous experts, including the members of the Steering Group Committee, Jack Triplett (consultant to the project team), Sir Tony Atkinson, Barbara Fraumeni, Andrew Jackson, Azim Lakhani, Phillip Lee, Alan Maynard, Alistair McGuire and Alan Williams. A number of individuals and organisations contributed to assemble the data used in this report. We are grateful to Kate Byram, Mike Fleming, Geoff Hardman, James Hemingway, Sue Hennessy, Sue Macran, Paula Monteith, Casey Quinn, Sarah Scobie, Bryn Shorney, Craig Spence, Karen Wagner, Chris Watson, BUPA, the Cardiff Research Consortium, Health Outcomes Group and York Hospitals NHS Trust. Any errors and omissions remain the sole responsibility of the authors.

References

Atkinson, T. (2005) 'Measurement of government output and productivity for the national accounts', Atkinson Review: Final Report, HMSO, 31 January 2005.

Baily, M. and Garber, A. (1997). 'Health care productivity', *Brookings Papers on Economic Activity. Microeconomics*, 143-215.

Baker, D., Einstadter, D., Thomas, C., Husak, S., Gordon, N., Cebul, R. (2002). 'Mortality trends during a program that publicly reported hospital performance', *Medical Car,e* 40 (10): 879-890

Basu, S., Fernald, J., Oulton, N. and Srinivasan, S. (2003) 'The case of the missing productivity, or, does information technology explain why productivity accelerated in the United States but not in the United Kingdom', *NBER working paper* no. 10010.

Berndt, E.R., Busch, S.H. and Frank, R.G. (2001) 'Treatment price indexes for acute phase major depression', in D.M. Cutler and E.R. Berndt, eds.: *Medical Care Output and Productivity*, University of Chicago Press, Chicago.

Berndt, E.R., Bir, A., Busch, S.H., Frank, R.G. and Normand, S.T. (2002) 'The medical treatment of depression, 1991-1996: productive inefficiency, expected outcome variations, and price indexes', *Journal of Health Economics*, 21: 373-396.

Bruce, J., Russell, E.M., Mollison, J. and Krukowski, Z.H. (2001) 'The measurement and monitoring of surgical adverse events', *Health Technology Assessment*, 5 (22). Available at: http://www.hta.nhsweb.nhs.uk/fullmono/mon522.pdf

Cairns, J.A. (1994) 'Valuing future benefits', *Health Economics,* 3: 221-229.

Cairns, J.A. and van der Pol, M.M. (1997) 'Constant and decreasing timing aversion for saving lives', *Social Science and Medicine,* 45 (11): 1653-1659.

Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N. and Spencer, A. (1999) 'The contingent valuation of safety and the safety of contingent valuation, part 2: The CV/SG 'chained' approach', *Journal of Risk and Uncertainty,* 17: 187-213.

CDRweekly (2002) The first year of the Department of Health's mandatory MRSA bacteraemia surveillance scheme in acute NHS Trusts I England: April 2001 – March 2002, Health Protection Agency, Vol. 12, nr. 25, 20 June 2002

CDRweekly (2003) The second year of the Department of Health's mandatory MRSA bacteraemia surveillance scheme in acute NHS Trusts I England: April 2002 – March 2003, Health Protection Agency, Vol. 13, nr. 25, 19 June 2003.

CDRweekly (2004) The third year of regional and national analysis of the Department of Health's mandatory MRSA bacteraemia surveillance scheme in acute NHS

Trusts I England: April 2001 – March 2004, Health Protection Agency, Vol. 14, nr. 29, 15 July 2004.

Commission for Health Improvement (2003) *NHS performance ratings acute trusts, specialist trusts, ambulance trusts 2002/2003*, Commission for Health Improvement: London. http://www.chi.nhs.uk/ratings/

Crowcroft, N. S. and M. Catchpole (2002) 'Mortality from methicillin resistant Staphylococcus aureus in England and Wales: analysis of death certificates', *British Medical Journal*, 325: 1390-1391.

Cutler, D.M. and Huckman, R.S. (2003) 'Technological development and medical productivity: the diffusion of angioplasty in New York state', *Journal of Health Economics*, 22: 187-217.

Cutler, D.M., McClellan, M., Newhouse, J.P. and Remler, D. (2001) 'Pricing Heart Attack Treatments', in D.M. Cutler and E.R. Berndt, eds.: *Medical Care Output and Productivity*, University of Chicago Press, Chicago.

Dawson, D., Gravelle, H., Kind, P., O'Mahony, M., Street, A. and Weale, M. (2004a) 'Developing new approaches to measuring NHS outputs and productivity', First Interim Report to Department of Health, CHE Technical Paper 31, July 2004. Available at: http://www.york.ac.uk/inst/che/tech.htm

Dawson, D., Gravelle, H., Kind, P., O'Mahony, M., Street, A. and Weale, M. (2004b) 'Measurement of NHS outputs and productivity growth: memorandum to Department of Health on data requirements', 30 September 2004.

Dawson, D., Gravelle, H., Kind, P., O'Mahony, M., Street, A. and Weale, M. (2004c) 'Developing new approaches to measuring NHS outputs and productivity: Data for productivity estimates', Second Interim Report to Department of Health, November 2004.

Deaton, A. and Muellbauer, J. (1980) *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge.

Department of Health (1996) 'Policy Appraisal and Health: a guide from the Department of Health', London GO70/38 3901.

Department of Health (2001) *NHS Performance Ratings: Acute Trusts 2000/01*, Department of Health: London. http://www.doh.gov.uk/performanceratings/2001/index.html

Department of Health (2002a) 'Reforming NHS Financial Flows: introducing payment by results', Department of Health: London.

Department of Health (2002b) 'Getting ahead of the Curve: a strategy for combating infectious diseases (including other aspects of health protection)', Department of Health: London. Available at: http://www.dh.gov.uk/assetRoot/04/06/08/75/04060875.pdf

Department of Health (2002c) *NHS Performance Ratings and Indicators: Acute*

Trusts, Specialist Trusts, Ambulance Trusts, Mental Health Trusts 2001/02, Department of Health: London. http://www.doh.gov.uk/performanceratings/2002/index.html

Department of Health (2003a) *Investing in General Practice, the new General Medical Services Contract*, Department of Health: London. http://www.dh.gov.uk/assetRoot/04/07/86/58/04078658.pdf

Department of Health (2003b) 'Surveillance of Healthcare Associated Infections', Publication and Statistics, Letters and Circulars, Department of Health: London. Available at: http://www.dh.gov.uk/assetRoot/04/01/34/10/04013410.pdf

Department of Health (2004a), 'The 'Experimental' NHS Cost Efficiency Growth measure', Department of Health: London.

Department of Health (2004b), *Departmental Report*, Department of Health: London.

Department of Health (2004c) 'Chief Executive's Report to the NHS', Department of Health: London.

Department of Health (2004d), Bloodborne MRSA infection rates to be halved by 2008 – Reid, Publication and Statistics, Press Release, Department of Health: London. Available at: http://www.dh.gov.uk/PublicationsAndStatistics/PressReleases/PressReleasesNotices/fs/en?CONTENT_ID=4093533&chk=MY%2BkD/

Department of Health (2004e), Infection control training for over one million NHS staff, Publication and Statistics, Press Release, Department of Health: London. Available at: http://www.dh.gov.uk/PublicationsAndStatistics/PressReleases/PressReleasesNotices/fs/en?CONTENT_ID=4093432&chk=7bcARL

Department of Health (2004f), MRSA surveillance system – results, Health Protection Agency Communicable Disease Surveillance Centre for the Department of Health, Department of Health: London. Available at: http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsStatistics/PublicationsStatisticsArticle/fs/en?CONTENT_ID=4085951&chk=HBt2QD

Department of Health (2005) CMO Update, 'New reporting system to improve patient safety', March.

Devlin, N. and Parkin, D. (2004) 'Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis', *Health Economics,* 13:437-452.

EuroQol Group (1990) 'EuroQol – a new facility for the measurement of health-related quality of life', *Health Policy*, 16: 199-208.

Eurostat (2001) *Handbook on price and volume measures in national accounts*. Luxembourg, Office for Official Publications of the European Communities.

Giuffrida A, Gravelle H, Roland M. (1999) 'Measuring Quality of Care with Routine Data: Avoiding Confusion between Performance Indicators and Health Outcomes', *British Medical Journal,* 319:94-98.

Gravelle, H. and Smith, D. (2001) 'Discounting for Health Effects in Cost-Benefit and Cost-Effectiveness Analysis', *Health Economics,* 10:587-599.

Healthcare Commission (2004) *2004 performance ratings*, Healthcare Commission: London. http://ratings2004.healthcarecommission.org.uk/

Healthcare Commission (2005) *2005 performance ratings*, Healthcare Commission: London. http://ratings2005.healthcarecommission.org.uk/

Health Protection Agency (2003), NINSS Partnership, Surveillance of Surgical Site Infection in English hospitals 1997 – 2002. Available at: http://www.hpa.org.uk/infections/topics_az/hai/SSIreport.pdf

Health Protection Agency and Office for National Statistics (2004), Trends in MRSA in England and Wales: analysis of morbidity and mortality data for 1993 – 2002, Health Statistics Quarterly 21 – Spring 2004. Available at: http://www.statistics.gov.uk/STATBASE/Product.asp?vlnk=6725

Health Protection Agency (2004), Protocol for Surveillance of Surgical Site Infection – Surgical Site infection Surveillance Service England. http://www.hpa.org.uk/infections/topics_az/hai/SSI%20Protocol.pdf

Hemingway, J. (2004) 'Sources and methods for public service productivity: health', *Economic Trends*, 613: 82-90.

Hicks, J.R. (1940) 'The valuation of social income', *Economica*, 7: 105-124.

HM Treasury (2003) 'The Green Book: Appraisal and Evaluation in Central Government', London: TSO.

Hofer, T.P. and Hayward, R.A. (1995) 'Can early re-admission rates accurately detect poor-quality hospitals?', *Medical Care,* 33(3): 234-45.

Hurst, J. and Siciliani, L. (2003) 'Tackling Excessive Waiting Times for Elective Surgery: A Comparison of Policies in Twelve OECD Countries', *OECD Health Working Paper 6*, OECD, Paris. Available at: www.oecd.org/health

Jackson G, Tobias M. (2001) 'Potentially avoidable hospitalizations in New Zealand, 1989-98', *Australian and New Zealand Journal of Public Health*, 25(3):212-21.

Jorgenson D, Ho. M and Stiroh, K. (2005) 'Labour Input and the Returns to Education', Chapter 6 of *Information Technology and the American Growth Resurgence*, MIT Press.

Lakhani, A., Coles, J., Eayres, D., Spence, C. & Rachet, B. (2005) 'Creative use of existing clinical and health outcomes data to assess NHS performance in England: Part 1-performance indicators closely linked to clinical care', *British*

Medical Journal, 330:1426-1431.

Lancaster, K. (1971) *Consumer Demand: A New Approach*, Columbia University Press, New York.

Lee, P. (2004) 'Public service productivity: health', *Economic Trends*, 613: 38-59.

Ludke, R.L., Booth, B.M. and Lewis-Beck, J.A. (1993) 'Relationship between early readmission and hospital quality of care indicators', *Inquiry,* 30: 95-103.

Mai, N. (2004) 'Measuring health care output in the UK: a diagnosis based approach.'

National Patient Safety Agency (2005) Building a memory: preventing harm, reducing risks and improving patient safety, July, on line resource available at http://www.npsa.nhs.uk/site/media/documents/1246_PSO_Report_FINAL.pdf

National Patient Safety Agency (2005) National Patient Safety Agency: helping to make the NHS safer, July 2005, on line resource available at: http://www.npsa.nhs.uk/site/media/documents/1260_NPSA_Information.pdf

OECD (2000*) A System of Health Accounts. Version 1.0*. OECD, Paris.

OECD (2001) *Measuring Productivity: Measurement of Aggregate and Industry Level Productivity Growth*, OECD, Paris. Available at: http://www.oecd.org/dataoecd/59/29/2352458.pdf

Pratt, J. (1964) 'Risk aversion in the small and the large', *Econometrica*, 32: 122-136.

Roland, M. (2004) 'Linking physician pay to quality of care – a major experiment in the UK', *New England Journal of Medicine*. 351, 1448-1454.

Rosen, S. (2002) 'Markets and diversity', *American Economic Review*, 92 (1): 1-15.

Ryan, M., Odejar, M. and Napper, M. (2004) 'The Value of Reducing Waiting Time in the Provision of Health Care: A Review of the Evidence.' Report to the Department of Health, Health Economics Research Unit, Aberdeen.

Sefton, J. and Weale, M. (to appear) 'The concept of income in a general equilibrium', *Review of Economic Studies*.

Shapiro, I., Shapiro, M.D. and Wilcox, D.W. (2001) 'Measuring the Value of Cataract Surgery', in D.M. Cutler and E.R. Berndt, eds.: *Medical Care Output and Productivity*, University of Chicago Press, Chicago.

Simkins, A. (2005) 'Using the GP contract Quality and Outcomes Framework for quality adjustment', mimeo, August 2005.

Street, A. and AbdulHassain, S. (2004), 'Would Roman soldiers fight for the financial flows regime? The re-issue of Diocletian's edict in the NHS', *Public Money and Management*, 24 (5): 301-38.

Thomas, J.W. (1996) 'Does risk-adjusted readmission rate provide valid information

on hospital quality?' *Inquiry* 1996; 28: 258-70.

Victorian Government (2004), The Victorian Ambulatory Care Sensitive Conditions Study, 2001–02, Public Health, Rural and Regional Health and Aged Care Services, Victoria Government Department of Human Services Melbourne Victoria, July 2004.

Williams, A. (1985) 'The economics of coronary artery bypass grafting', *British Medical Journal*, 291: 326-9.

Yule, G.U. (1934) 'On some points relating to vital statistics, more especially statistics of occupational mortality', *Journal of the Royal Statistical Society*, 97: 1-84.

Annex: How should NHS output be measured?

Value weighted output index

The *value weighted output index* is our preferred way to measure NHS output:

$$I_{yt}^{xq} = \frac{\sum_j x_{jt+1} \sum_k \pi_{kt} q_{kjt+1}}{\sum_j x_{jt} \sum_k \pi_{kt} q_{kjt}}$$

where x_{jt} is the volume of output j in period t, q_{kjt} is the amount of outcome or characteristic k produced by a unit of j, and π_{kt} is marginal value of outcome k.[11]

The index requires data on both the characteristics produced and on their marginal social value. Since improving the health of patients is a primary objective of the NHS, improved health outcomes are one of the most important characteristics of treatment. But other characteristics of treatment also affect utility, e.g. the length of time waited for treatment, the degree of uncertainty attached to the waiting time, relationship with doctors, hospital food and safety. These can be incorporated in the value weighted output index when the necessary data are available.

Cost weighted output index

Continuous inpatient spells (CIPS)

If the data needed to calculate the value weighted output index are not available, we can instead use unit costs to weight outputs, and make use of available data to quality adjust these cost weighted outputs. The *cost weighted output index* is:

$$I_{ct}^{x} = \frac{\sum_j x_{jt+1} c_{jt}}{\sum_j x_{jt} c_{jt}}$$

Information on survival can be used to adjust the cost weighted output index

30 day mortality rates

Data on short term survival can be used to adjust the index as follows:

$$\frac{\sum_j c_{jt} x_{jt+1} \left(\frac{a_{jt+1}}{a_{jt}} \right)}{\sum_j c_{jt} x_{jt}}$$

Information on long term survival (not currently available for most treatments) could be used to adjust the index as follows:

$$\frac{\sum_j x_{jt+1} c_{jt} \frac{a_{jt+1}}{a_{jt}} \sum_{s=1}^{S} \left(\frac{\sigma_{jt+1}^{*}(s)}{\sigma_{jt}^{*}(s)} \right) \left(\frac{\sigma_{jt}^{*} \delta^{s}}{\sum_{s=1}^{5} \sigma_{jt}^{*}(s) \delta^{s}} \right)}{\sum_j x_{jt} c_{jt}}$$

[11] Please refer to the table of notation at the end of this report for further details of the notation used here.

The simple survival adjustment above implies that the patient would have zero quality adjusted life years if not treated. It is possible to introduce an additional term into the formula to include a uniform estimate of the difference between health before and after treatment, giving the *health effect survival index*:

$$\frac{\sum_j x_{jt+1}\left(\frac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)c_{jt}}{\sum_j x_{jt}c_{jt}}$$

Note that for HRGs where the mortality rate is high, we make no health effect adjustment and use only the change in survival.

Information on changes in the life expectancy of patients treated can also be included as follows:

$$\frac{\sum_j x_{jt+1}c_{jt}\frac{\left(a_{jt+1}-k_j\right)}{\left(a_{jt}-k_j\right)}\left(\frac{1-e^{-rL_{jt+1}}}{1-e^{-rL_{jt}}}\right)}{\sum_j x_{jt}c_{jt}}$$

Certainty equivalent wait – measured as the waiting time for patients at the 80th percentile of the waiting time

Information on waiting times can be used to quality adjust the cost weighted output index. This approach regards reductions in the wait for treatment as valuable because of their effect on the discounted value of the health gain from treatment

There are two main forms of waiting time adjustment

Discount to date of treatment with charge for waiting

$$\frac{\sum_j x_{jt+1}c_{jt}\left(\frac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)\frac{\left[\frac{\left(1-e^{-r_L L_{jt+1}}\right)}{r_L}-\frac{\left(e^{r_w w_{jt+1}}-1\right)}{r_w}\right]}{\left[\frac{\left(1-e^{-r_L L_{jt}}\right)}{r_L}-\frac{\left(e^{r_w w_{jt}}-1\right)}{r_w}\right]}}{\sum_j x_{jt}c_{jt}}$$

This is the form of the CWOI that we recommend should be used in the interim, where;

- $r_L = r_W = 0.015$,
- k_j = actual k_j if known, = mean k for known electives if k_j not known and elective, = ½ mean k for electives if non-elective,
  - if $(a_{jt+1} - k_j)$ and $(a_{jt} - k_j) < 0.10$ then $k_j = 0$.
  - w_{jt}, w_{jt+1} are 80th percentile waits

Discount to date placed on list

$$\frac{\sum_j c_{jt}x_{jt+1}\left(\frac{a_{jt+1}-k_j}{a_{jt}-k_j}\right)\left(\frac{e^{-rw_{jt+1}}\left(1-e^{-rL_{jt+1}}\right)}{e^{-rw_{jt}}\left(1-e^{-rL_{jt}}\right)}\right)}{\sum_j c_{jt}x_{jt}}$$

(assumes $r_W = r_L = r$)

Additional quality adjustments can also be made to the CWOI. A lack of appropriate data means that only illustrative adjustments for these additional aspects of quality could be calculated at present.

A quality adjustment for readmissions and MRSA can be incorporated based on the assumption that their cost is a deadweight loss which reduces the value of treatment. Hence the CWOI (ignoring other quality adjustments for illustrative purposes) can be adjusted as follows,

$$\frac{\sum_j x_{jt+1}c_{jt} - \sum_j x^b_{jt+1}c^b_{jt}}{\sum_j x_{jt}c_{jt} - \sum_j x^b_{jt}c^b_{jt}}$$

where x^b denotes the number of readmissions or cases of MRSA and c^b their costs.

Patient satisfaction can also be incorporated in the CWOI, using measures of patient experience, derived from patient satisfaction surveys, as summary indicators of characteristics that patients value.
This can be incorporated as follows;

$$I_t^{Comp} = \sum_{k=0}^{n} \omega_k I_t^k$$

where I_t^{Comp} is calculated as the weighted sum of the growth rate of one of the quality adjusted output indices and the growth rates of the other indicators. We denote the growth rate of indicator k by I_t^k. If there are n such indicators, and the relevant weights are denoted by ω_k then the overall index is given by the formula above.

Table of Notation

Notation	Interpretation
x_{jt}	quantity of output *j* at time *t* (units of *j*)
π_{kt}	marginal social value of characteristic *k* at time *t*
q_{jkt}	quantity of characteristic *k* produced by one unit of output *j* at time *t*
$p_{jt} = \sum_k \pi_{kt} q_{jkt}$	marginal social value of unit of output *j* at time *t* (£s per unit of *j*)
$y_{jt} = p_{jt} x_{jt}$	value of output *j* at time *t* (£s)
$y_t = \sum_i y_{jt}$	total value of NHS output
I_{yt}^{xq}	value weighted output index
g_{xjt}	growth rate of output x_j
g_{qkjt}	growth rate of characteristic *k* produced by output *j*
ω_{pt}^{kt}	proportion of marginal value of output *j* accounted for by characteristic *k*
ω_{yt}^{jt}	proportion of the total value of period *t* output accounted for by output *j*
I_{ct}^{x}	cost weighted output index CWOI
c_{jt}	unit (average) cost of output *j* at time *t* (£s per unit of *j*)
m_{jt}	mortality rate from NHS output *j* in period *t*
a_{jt}	survival rate (1-m_{jt})
θ	vector of mental and physical health characteristics
$h(\theta)$	health level from having health state θ
$\sigma_{jt}^{*}(s)$	probability of surviving *s* periods given that the patient survived treatment *j* at date *t*
$p_{jt}^{*}(\theta, s)$	probability of being in health state θ conditional on surviving s periods after treatment *j* at date *t*
$h_{jt}^{*}(s) = \sum_\theta p_{jt}^{*}(\theta, s) h(\theta)$	expected level of health conditional on surviving *s* periods
δ	discount factor
$q_{jt}^{*} = \sum_s \delta^{s-t} \sigma_{jt}^{*}(s) h_{jt}^{*}(s)$	discounted sum of quality adjusted life years produced by the treatment if the patient survives treatment
$h_j^{o}(s)$	expected health *s* periods hence if

	the patient does not receive treatment *j* conditional on surviving *s* periods
$\sigma_j^0(s)$	probability of surviving without treatment
$p_j^0(\theta, s)$	probability of being in health state θ after *s* periods conditional on surviving without receiving treatment
$q_{jt}^0 = \sum_s \delta^s \sigma_j^o(s) h_j^o(s)$	discounted sum of quality adjusted life years if patient not treated
$q_{jt} = (1 - m_{jt}) q_{jt}^* - q_{jt}^0$	expected increase in discounted QALYs from treatment *j* at time *t*
g_{ajt}	growth rate of survival
$g_{q_{jt}^*}$	growth rate in q_{jt}^*
$g_{q_{jt}^0}$	growth rate in q_{jt}^0
I_{ct}^{xa}	survival adjusted cost weighted output index
π_{wjt}	value of a reduction of one day in waiting time for treatment *j* in year *t*
π_{ht}	value of health gain
r_w, r_L	discount rates on the wait for treatment, QALYs
w_{jt}	waiting time for treatment *j* in year *t*
$h^*(s; w_t)$	time path of health with treatment after wait *w*
L_t^0	life expectancy without treatment
L_t^*	life expectancy with treatment
I_{ct}^{xaw}	survival and waiting time adjusted cost weighted index